



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 134616**

**TO: James Schultz**  
**Location: rem/2d18/2c18**  
**Art Unit: 1635**  
**Thursday, October 07, 2004**

**Case Serial Number: 10/054313**

**From: Alex Waclawiw**  
**Location: Biotech-Chem Library**  
**Rem 1A71**  
**Phone: 272-2534**

**Alexandra.waclawiw@uspto.gov**

### **Search Notes**

This Page Blank (Empty)

STIC-Biotech/ChemLib

CRFE

134616

mej

From: Schultz, James  
Sent: Wednesday, October 06, 2004 5:21 PM  
To: STIC-Biotech/ChemLib  
Subject: Seq search 10/054,313

Hello,  
Please run a standard amino acid sequence search on SEQ ID NO:1, including the interference databases.

Thanks  
Doug Schultz

James Douglas Schultz, PhD  
AU 1635 (Biotechnology)  
Patent Examiner  
United States Patent and Trademark Office  
(Office) REM 2D18  
(Mail) REM 2C18  
(571) 272-0763

1-AA-286

mej

RECEIVED  
OCT - 7 2004  
TECH/CHEM. DIVISION  
(STIC)

6

Point of Contact:

Alexandra Wacławiw

Technical Info. Specialist

STAFF USE ONLY  
CM 16A02 Tel: 308-4491

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 10-7-04  
Date Completed: 10-7-04  
Searcher Prep/Rev. Time: 6  
Online Time: 7

\*\*\*\*\*

Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # (C) \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): CompuLink

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:02:55 / Search time 117 Seconds

(without alignments)  
771.267 Million cell updates/sec

Title: US-10-054-313-1

Perfect score: 1546  
Sequence: 1 MSWLFALHRAALALPCRR.....FIGNEADRLAREGKQSD 286

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 100 summaries

SPREMB-25+\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_ricent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioph:\*  
17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	1193.5	77.2	285	11	Q8VCR6	Q8VCR6 mus musculus
2	882	57.1	293	13	Q91953	Q91953 gallus galli
3	744	48.1	153	11	Q8BU23	Q8BU23 mus musculus
4	491	31.8	251	5	Q86MG7	Q86MG7 caenorhabdi
5	422	27.3	333	5	Q9Y335	Q9Y335 drosophila
6	422	27.3	333	5	Q9Y335	Q9Y335 drosophila
7	370.5	24.0	301	5	Q00870	Q00870 trypanosoma
8	323.5	20.9	264	3	Q9UST8	Q9UST8 schizosacch
9	245	15.8	159	16	Q8DM24	Q8DM24 synecococc
10	241.5	15.6	154	16	Q89UJ3	Q89UJ3 bradyrhizob
11	241	15.4	298	16	Q88UB4	Q88UB4 lactobacilli
12	238	15.4	302	16	Q8Z0F5	Q8Z0F5 anabaena sp
13	221.5	14.3	153	16	Q7VM15	Q7VM15 haemophilus
14	221.5	14.3	161	16	Q82XV8	Q82XV8 nitrosomona
15	219.5	14.2	149	5	Q81T47	Q81T47 leishmania
16	219.5	14.2	150	16	Q87YT0	Q87YT0 pseudomonas

17	218.5	14.1	148	16	Q88FF5	Q88FF5 pseudomonas
18	214	13.8	202	16	Q934V8	Q934V8 salmonella
19	209.5	13.6	155	16	Q7WC18	Q7WC18 bordetella
20	209.5	13.6	155	16	Q7W0T2	Q7W0T2 bordetella
21	209	13.5	167	16	Q7VQB6	Q7VQB6 candidatus
22	208.5	13.5	257	16	Q7V474	Q7V474 prochloroc
23	208.5	13.5	155	16	Q7VRX8	Q7VRX8 bordetella
24	208	13.5	248	16	Q7UJ31	Q7UJ31 synecococc
25	203.5	13.2	158	16	Q8EE30	Q8EE30 shewanella
26	203.5	13.2	169	16	Q92S39	Q92S39 rhizobium m
27	203	13.1	328	16	Q8G3X8	Q8G3X8 bifidobacte
28	200	12.9	140	16	Q7V385	Q7V385 prochloroc
29	196.5	12.7	154	16	Q83EK3	Q83EK3 coxiella bu
30	196	12.7	165	16	Q7VDY9	Q7VDY9 prochloroc
31	195	12.6	165	16	Q87RY4	Q87RY4 vibrio para
32	194.5	12.6	157	2	Q8RTZ8	Q8RTZ8 shewanella
33	194	12.5	169	16	Q87U15	Q87U15 vibrio para
34	193	12.5	240	16	Q97G21	Q97G21 clostridium
35	191.5	12.4	139	5	Q9XVE6	Q9XVE6 caenorhabdi
36	184	11.9	418	5	Q86LS8	Q86LS8 caenorhabdi
37	184	11.9	477	5	Q09633	Q09633 caenorhabdi
38	182	11.8	158	16	Q7UF86	Q7UF86 rhodospirae
39	180	11.6	146	16	Q93SU7	Q93SU7 chlorobium
40	178	11.5	206	16	Q892E9	Q892E9 clostridium
41	176.5	11.4	236	16	Q93HL3	Q93HL3 streptomyc
42	170	11.0	209	16	Q8XKP0	Q8XKP0 clostridium
43	169.5	11.0	146	16	Q8KUL7	Q8KUL7 vibrio chol
44	169	10.9	220	16	Q8EU10	Q8EU10 mycoplasma
45	169	10.9	482	9	Q8SD07	Q8SD07 pseudomonas
46	164	10.6	1127	15	Q70652	Q70652 gibbon ape
47	158.5	10.3	1786	15	Q89811	Q89811 feline leuk
48	158	10.2	223	16	Q9X122	Q9X122 thermotoga
49	157	10.2	179	16	Q9RXV2	Q9RXV2 deinococcus
50	157	10.2	1145	6	Q9XSN8	Q9XSN8 sus scrofa
51	156.5	10.1	161	16	Q83HK9	Q83HK9 tropheryma
52	156.5	10.1	477	15	Q83GM3	Q83GM3 tropheryma
53	155	10.0	167	16	Q85732	Q85732 woolly monk
54	155	10.0	1738	15	Q93735	Q93735 feline muri
55	155	10.0	1784	15	Q85521	Q85521 feline leuk
56	154	10.0	1127	6	Q9TIC1	Q9TIC1 phascolatart
57	153	9.9	1193	15	Q90RL9	Q90RL9 porcine end
58	152	9.8	300	15	Q11980	Q11980 murine leuk
59	152	9.8	868	15	Q73505	Q73505 porcine end
60	152	9.8	1144	15	Q8UM99	Q8UM99 porcine end
61	152	9.8	1146	15	Q8UM95	Q8UM95 porcine end
62	152	9.8	1204	15	Q41250	Q41250 raucher mu
63	152	9.8	1718	15	Q8J4V8	Q8J4V8 porcine end
64	152	9.8	1720	15	Q8J4V6	Q8J4V6 porcine end
65	152	9.8	1733	15	Q9E7M1	Q9E7M1 dg-75 murin
66	152	9.8	2376	15	Q9Q1X3	Q9Q1X3 porcine end
67	152	9.8	2376	15	Q9Q1X5	Q9Q1X5 porcine end
68	152	9.8	2378	15	Q9Q1X4	Q9Q1X4 porcine end
69	150	9.7	397	15	Q83497	Q83497 murine leuk
70	150	9.7	1737	15	Q92808	Q92808 moloney mur
71	149	9.6	669	15	Q9E0U7	Q9E0U7 murine leuk
72	149	9.6	1195	15	Q90RL6	Q90RL6 porcine end
73	149	9.6	1204	15	Q9YK99	Q9YK99 murine leuk
74	149	9.6	1736	15	Q7ZUT6	Q7ZUT6 amphotropic
75	146.5	9.5	1139	15	Q806U7	Q806U7 porcine end
76	146	9.4	921	11	Q80T01	Q80T01 mus musculus
77	146	9.4	1146	15	Q8UM96	Q8UM96 porcine end
78	146	9.4	1204	15	Q7ZKX7	Q7ZKX7 recombinant
79	146	9.4	1738	15	Q8UN00	Q8UN00 moloney mur
80	146	9.4	2237	13	Q8U1T1	Q8U1T1 brachydanto
81	145.5	9.3	1117	2	Q54388	Q54388 streptomyc
82	144.5	9.3	1043	15	Q90DD2	Q90DD2 chimpanzee
83	144	9.3	1203	11	Q89815	Q89815 mus musculi
84	143	9.2	707	15	Q61918	Q61918 murine leuk
85	143	9.2	1038	15	Q83398	Q83398 murine leuk
86	143	9.2	1734	15	Q9J8E2	Q9J8E2 murine leuk
87	143	9.2	1734	15	Q7SVK7	Q7SVK7 murine leuk
88	142.5	9.2	1124	15	Q66972	Q66972 feline immu
89	142	9.2	672	15	Q83379	Q83379 rat leukemi

90 140 9.1 1142 15 Q8Q6U4 Q8Q6U4 porcine end  
 91 140 9.1 1196 15 Q90RL4 Q90RL4 murine leuk  
 92 139 9.0 1734 11 P70355 P70355 mus musculus  
 93 138 8.9 133 2 Q54222 Q54222 streptomyces  
 94 138 8.9 133 2 Q54222 Q54222 streptomyces  
 95 138 8.9 133 2 Q54222 Q54222 streptomyces  
 96 136 8.9 1736 15 Q83362 Q83362 murine leuk  
 97 135.5 8.8 1736 15 Q83362 Q83362 murine leuk  
 98 135.5 8.8 1736 15 Q83362 Q83362 murine leuk  
 99 134.5 8.7 1059 15 Q76630 Q76630 human immun  
 100 134.5 8.7 1432 15 Q82851 Q82851 jembrana di

## ALIGNMENTS

RESULT 1  
 Q8VCR6 PRELIMINARY; PRT; 285 AA.  
 ID Q8VCR6; PRELIMINARY; PRT; 285 AA.  
 AC Q8VCR6; PRELIMINARY; PRT; 285 AA.  
 DT 01-MAR-2002 (TEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Ribonuclease H1.  
 DE Ribonuclease H1.  
 GN RNASEH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC019411; AA019411.1; -.  
 DR MGD: MGI:1335073; Rnaseh1.  
 DR GO: GO:0003676; F:nucleic acid binding; IEA.  
 DR GO: GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro: IPR009027; L9 N like.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 SQ SEQUENCE 285 AA; 31833 MW; 843482077C7E9230 CRC64;  
 Query Match 77.2%; Score 1193.5; DB 11; Length 285;  
 Best Local Similarity 76.9%; Pred. No. 8.7e-103;  
 Matches 220; Conservative 27; Mismatches 38; Indels 1; Gaps 1;  
 QY 1 MSWLLFLARHVALAALPCRRGSGFGFVAVRRGRTGVFLTNNEGRAQVDRPPARFKK 60  
 DB 1 MWLLPLSRVTTLAVVRLRRGICGLGIFAVRRGRRTGVFLSWSECKAQVDRPPARFKK 60  
 QY 61 FATEDAMAFVRRKSPVSEVSEGHENHGHSEAKPGKRLREPLDGDGHEGAQVYAKHMKP 120  
 DB 61 FATEDAMAFVRRKSPVSEVSEGHENHGHSEAKPGKRLREPLDGDGHEGAQVYAKHMKP 119  
 QY 121 SVEPAPVSRDTFSYMGDFVAVVYTDGCCSNGRRKPRAGIGVYWGPHPLVNGIRLRGRQ 180  
 DB 120 DTEPAVVSVDTSYMGESVIVYTDGCCSNGRRKPRAGIGVYWGPHPLVNGIRLRGRQ 179  
 QY 181 TNGRAETHAACKIAEQAKTONINKLYTDSMTINGITWVQGGKNGKTSAGKEVIN 240  
 DB 180 TNGRAETHAACKIAEQAKTONINKLYTDSMTINGITWVQGGKNGKTSAGKEVIN 239  
 QY 241 KEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKOSED 286  
 DB 240 KEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKOSED 285

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE mRNA, complete cds, clone CUFEST65.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxId=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=lens fibers;  
 RA MEDLINE=96437509; PubMed=8440185;  
 RA Sawada K., Agata K., Eguchi G.;  
 RT "Characterization of terminally differentiated cell state by  
 RT categorizing cDNA clones derived from chicken lens fibers."  
 RL Int. J. Dev. Biol. 40:531-535(1996).  
 DR EMBL: D26340; BA05382.1; -.  
 DR HSSP: Q04740; IQHK.  
 DR GO: GO:0003676; F:nucleic acid binding; IEA.  
 DR GO: GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro: IPR009027; L9 N like.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 SQ SEQUENCE 293 AA; 32499 MW; 2B86AD6DC722682B CRC64;  
 Query Match 57.1%; Score 882; DB 13; Length 293;  
 Best Local Similarity 57.7%; Pred. No. 9.2e-74;  
 Matches 173; Conservative 34; Mismatches 69; Indels 24; Gaps 5;

QY 1 MSWLLFLARHVALAALPCRRGSGFGFVAVRRGRTGVFLTNNEGRAQVDRPPARFKK 60  
 DB 2 LRLW-----VALLSHSC-FVSKGGMFVAVRRGRGTGVRTVAECQQOVNRRPSPASFKK 54  
 QY 61 FATEDAMAFVRRKSPVSEVSEGHENHGHSEAKPGKRLREPLDGDGHEGAQVYAKHMKP 114  
 DB 55 FATEDAMAFVRRKSPVSEVSEGHENHGHSEAKPGKRLREPLDGDGHEGAQVYAKHMKP 114  
 QY 115 -----AKHMKSPVSEVSEGHENHGHSEAKPGKRLREPLDGDGHEGAQVYAKHMKP 165  
 DB 115 TNEHTVRAKHI--DEQSTPVVSEAKFSYMGESVIVYTDGCCSNGRRKPRAGIGVYWG 172  
 QY 166 RGHPLVNGIRLRLPORNORAEIHAACKIAEQAKTONINKLYTDSMTINGITWVQGGK 225  
 DB 173 RGHPLVNGIRLRLPORNORAEIHAACKIAEQAKTONINKLYTDSMTINGITWVQGGK 232  
 QY 226 KKGKMTSAGKEVYKEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKOSED 285  
 DB 233 KKGKMTSAGKEVYKEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKOSED 292  
 RESULT 3  
 Q8BU23 PRELIMINARY; PRT; 153 AA.  
 ID Q8BU23; PRELIMINARY; PRT; 153 AA.  
 AC Q8BU23; PRELIMINARY; PRT; 153 AA.  
 DT 01-MAR-2003 (TEMBLrel. 23, Created)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Similar to ribonuclease H1.  
 GN RNASEH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium;  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT Nature 420:563-573(2002)."  
 RL Nature 420:563-573(2002).  
 DR EMBL: AK088032; BAC40109.1; -.

DR MGD; MGI:1335073; Rnaaseh1.  
DR GO; GO:0003676; F:ribonuclease H binding; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR InterPro; IPR002156; Rnaaseh.  
DR Pfam; PF00075; rnaaseh; 1.  
DR SEQUENCE 153 AA; 16922 MW; C9AA7E094274F12E CRC64;  
SQ  
Query Match 48.1%; Score 744; DB 11; Length 153;  
Best Local Similarity 87.4%; Pred. No. 2.7e-61;  
Matches 132; Conservative 12; Mismatches 7; Indels 0; Gaps 0;  
QY 136 MGFVVVYTDGCCSSNGRRKPRAGIGYVWGPCHPLNVGIRLPGRQTNQRAEIIAAACKAIE 195  
DB 3 VGESIVYVTDGCCSSNGRRKPRAGIGYVWGPCHPLNVGIRLPGRQTNQRAEIIAAACKAIM 62  
QY 196 QAKTONINKLVLYTDSMTFINGITNWVGQWKKNGKTSAGKEVINKEDFVALERLTQGM 255  
DB 63 QAKQNIISKLVLYTDSMTFINGITNWVGQWKKNGKTSAGKEVINKEDFVALERLTQGM 122  
QY 256 IOMMHVPGHSGFTSGNEADRLAREGAKOSED 286  
DB 123 IOMMHVPGHSGFTSGNEADRLAREGAKOSED 153  
RESULT 4  
ID 086MG7 PRELIMINARY; PRT; 251 AA.  
AC 086MG7;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein F59A6.9.  
GN F59A6.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA MEDLINE=99069613; PubMed=9851916;  
RA Watson R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Nham M.;  
RT "The sequence of C. elegans cosmid F59A6";  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U41994; AA091712.1; -  
DR WormPep; F59A6.9; C533660.  
DR GO; GO:0003676; F:ribonuclease H binding; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR InterPro; IPR009027; L9 N like.  
DR InterPro; IPR002156; Rnaaseh.  
DR Pfam; PF00075; rnaaseh; 1.  
KM Hypothetical protein.  
SQ SEQUENCE 251 AA; 28017 MW; CFE6A1B5E0DF9A CRC64;  
Query Match 31.8%; Score 491; DB 5; Length 251;  
Best Local Similarity 39.6%; Pred. No. 1.9e-37;  
Matches 107; Conservative 39; Mismatches 78; Indels 46; Gaps 5;  
QY 28 FVAVRGRKTVPLTNECRAQVDRPAPAFKKPATDEAMAFV-----RKSASPV 79  
DB 7 YVAVARGQVGIYRTNECKTOIDGFONARFKKFAEARKFVADNMVSGSKPVTAV 66

QY 80 SEG-----HENOHGSEAKPGKRLREPLDGDGHSQAQPYAKRMKPSVEPAPPVSRDT 132  
DB 67 STSSATRKRTHEGT--KTEAKKMTBEEVD-----PEFANAP----- 103  
QY 133 FSYMGDPVVVYTDGCCSSNGRRKPRAGIGYVWGPCHPLNVGIRLPGRQTNQRAEIIAAACK 192  
DB 104 -----VYVTDGACSSNGTKNAKAGVYWGDDSEDNREGPYVGPATNNRGELIAVQK 155  
QY 193 AIEQAKTONINKLVLYTDSMTFINGITNWVGQWKKNGKTSAGKEVINKEDFVALERLTQ 252  
DB 156 AIEKAIERKLPVYVTKTDSNLLVQSMNVIHWGKKRKGKTSAGSEVYLVQVDMKIDNLRQ 215  
QY 253 GMDIOMMHVPGHSGFTSGNEADRLAREGAK 282  
DB 216 KLVKVFHVHVGAGIDGNEKADRLARKGAQ 245  
RESULT 5  
ID 09V335 PRELIMINARY; PRT; 333 AA.  
AC 09V335;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE RNH1 protein (AT19436p).  
GN RNH1 OR CG8729.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA George R.A., Lewis S.E., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA Amantides P.G., Scherer S.E., Yandell M.D., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zheng Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borzova D., Botchan M.R., Bouck U., Brocktein P., Broctier P.,  
RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keithum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA LaSro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier B., Spirdling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskae R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195(2000).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA George R., Gonzalez M., Guarin H., Dreesnek D., Farfan D., Fisse E.,  
 RA George R., Gonzalez M., Guarin H., Dreesnek D., Farfan D., Fisse E.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacible J., Paragas V., Park S.,  
 RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AEO03839; AAF59170.1; -  
 DR EMBL: AY089374; AAF59112.1; -  
 DR HSSP: Q04740; IQHK.  
 DR FlyBase: FBgn0023171; rnh1.  
 DR GO: GO:0004524; F:ribonuclease H1 activity; IDA.  
 DR InterPro: IPR009027; L9 N like.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaeH; 1.  
 SQ SEQUENCE 333 AA; 37148 MW; C52BEC83426B599D CRC64;

Query Match 27.3%; Score 422; DB 5; Length 333;  
 Best Local Similarity 34.4%; Pred. No. 7.5e-31;  
 Matches 107; Conservative 31; Mismatches 113; Indels 60; Gaps 7;

QY 28 FVAVRRGKRGVLTNECRQAYDRFPAPFKFATEDAMAFVR--KSASPE----- 78  
 DB 18 FVAVAGRRSGVYGSWAECEQYKFKKTRQADQFVNGCKSYAPDVAVPLG 77  
 QY 79 -----VSEGHENHG-----QSEAPK----- 95  
 DB 78 KEKASLASWKNSIEVKNPKRYTDEWPEEDHDLAEDDLNAAANEVEGDPKSNSSNLPDIL 137  
 QY 96 GKLRREPLDGDGHEAQAQYPAKHKPSEVAPRVSRDTFSYMCD---FVVVYTDGCCSSNG 152  
 DB 138 NRKRKGTSGDKXNKI PRHSAQVS---EATGLKQVAFQFEIDAEGVIVYTDGSCIGNG 194  
 QY 153 RRRPRAGIGYVWGPGHPLVNGIRLPGRTQNRALHAACALIQAKTQNIKLVLYTDSM 212  
 DB 195 RAGACAGYGYVFGKHNQNLNAKPEVGRVTNNVGEIQAAIHAITALDLGIQKLCISTDSQ 254  
 QY 213 FTINGITNNVQGWKKGWKTSAKEVINKEPFVALLRLTQ--GMDIQMHWVPGHSGFIGN 270  
 DB 255 FLINSTITLWAGKKGKDKMLKNNQPVKNVDFKELDKLQDNNITVKNVYEAHKGIEGN 314  
 QY 271 EADRLARLARGA 281  
 DB 315 EMADKLARQGS 325

RESULT 6  
 ID 044114 PRELIMINARY; PRT; 333 AA.  
 AC 044114;  
 DT 01-JUN-1998 (TEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Ribonuclease H1.  
 GN RNH1 OR CG8729.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Dirosophillidae; Drosophila.  
 OC NCB1\_TaxId=7227;  
 RX MEDLINE=98063328; PubMed=9398656;  
 RA Filippov V., Filippova M., Gill S.S.;  
 RT "Functional characterization of RNase H1 from Drosophila  
 RT melanogaster."  
 RL Biochem. Biophys. Res. Commun. 240:844-849(1997).  
 DR EMBL: AF032821; AAC74810.1; -.  
 DR PIR: JCS787; JCS787.

DR HSSP: Q04740; IQHK.  
 DR FlyBase: FBgn0023171; rnh1.  
 DR GO: GO:0004524; F:ribonuclease H1 activity; IDA.  
 DR InterPro: IPR009027; L9 N like.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaeH; 1.  
 SQ SEQUENCE 333 AA; 37164 MW; D180BC3436B5826 CRC64;

Query Match 27.3%; Score 422; DB 5; Length 333;  
 Best Local Similarity 34.4%; Pred. No. 7.5e-31;  
 Matches 107; Conservative 31; Mismatches 113; Indels 60; Gaps 7;

QY 28 FVAVRRGKRGVLTNECRQAYDRFPAPFKFATEDAMAFVR--KSASPE----- 78  
 DB 18 FVAVAGRRSGVYGSWAECEQYKFKKTRQADQFVNGCKSYAPDVAVPLG 77  
 QY 79 -----VSEGHENHG-----QSEAPK----- 95  
 DB 78 KEKASLASWKNSIEVKNPKRYTDEWPEEDHDLAEDDLNAAANEVEGDPKSNSSNLPDIL 137  
 QY 96 GKLRREPLDGDGHEAQAQYPAKHKPSEVAPRVSRDTFSYMCD---FVVVYTDGCCSSNG 152  
 DB 138 NRKRKGTSGDKXNKI PRHSAQVS---EATGLKQVAFQFEIDAEGVIVYTDGSCIGNG 194  
 QY 153 RRRPRAGIGYVWGPGHPLVNGIRLPGRTQNRALHAACALIQAKTQNIKLVLYTDSM 212  
 DB 195 RTGACAGYGYVFGKHNQNLNAKPEVGRVTNNVGEIQAAIHAITALDLGIQKLCISTDSQ 254  
 QY 213 FTINGITNNVQGWKKGWKTSAKEVINKEPFVALLRLTQ--GMDIQMHWVPGHSGFIGN 270  
 DB 255 FLINSTITLWAGKKGKDKMLKNNQPVKNVDFKELDKLQDNNITVKNVYEAHKGIEGN 314  
 QY 271 EADRLARLARGA 281  
 DB 315 EMADKLARQGS 325

RESULT 7  
 ID 000870 PRELIMINARY; PRT; 301 AA.  
 AC 000870;  
 DT 01-JUL-1997 (TEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Ribonuclease H1.  
 GN RNH1.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OC NCB1\_TaxId=5691;  
 RX MEDLINE=97321566; PubMed=9178276;  
 RA Hestlein D.G., Campbell A.G.;  
 RT "Molecular cloning and expression of a ribonuclease H from the  
 RT kinetoplastid, Trypanosoma brucei."  
 RL Mol. Biochem. Parasitol. 86:121-126(1997).  
 DR EMBL: U74470; AAC47537.1; -.  
 DR HSSP: Q04740; IQHK.  
 DR GO: GO:0003676; F:nucleic acid binding; IEA.  
 DR GO: GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro: IPR009027; L9 N like.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaeH; 1.  
 SQ SEQUENCE 301 AA; 33100 MW; C03984F7F3B75B3F CRC64;

Query Match 24.0%; Score 370.5; DB 5; Length 301;  
 Best Local Similarity 34.8%; Pred. No. 4.1e-26;  
 Matches 106; Conservative 33; Mismatches 103; Indels 63; Gaps 12;

QY 28 FVAVRRGKRGVLTNECRQAYDRFPAPFKFATEDAMAFVRKSASPEVSEGHENHG 87  
 DB 6 FVAVAGRTGVTWEECKQVSGFGSGARFKSFLTLQEHQAY--SGVQDVDPGSGVGTG 63

```

QY 88 GQSEAKP---GKRLREPLD-----GDGHSAGPVAKMKSPVAPVSRDTF 133
DB 64 LVADSHTEPLVAVKALSLDVLDAVEVDADGNE-VDDESKQVLTSEBWEARKREA- 121
QY 134 SYMGDFVVVYTDGSSNSG---RRKPRAGIGVYWGPGHPLNVGIRLPGR--QTNORAB- 186
DB 122 -----VVVVYVDGACGNNNSRSRERRPRAGFGFYGDGDSRNPKFPLPAHEPQTNORABL 175
QY 187 ---THACCAIEQAKTQINIKLVLTDSMFTINGITNNVQGWKKGWTKTSAGKEVINKE- 242
DB 176 SALIHVRLVADLSDHPCYN---LCVYSDSKYTWGVNSYLHWRERNGFKTAGGADVANDL 232
QY 243 ---DFVAL-----ERLT-----QGMIDQMHHVPGHSGFIGNEADRLA 277
DB 223 WSGFTKLRRHLSRCAERFTMEPRFKASLAATAARVALQLKHVPGHAGVYGNEMADRLA 292
QY 278 REGAK 282
DB 293 VEACE 297

RESULT 8
Q9UST8 PRELIMINARY; PRT; 264 AA.
ID Q9UST8 042798;
AC 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ribonuclease H (EC 3.1.26.4) (RNase H).
GN RNH1 OR SPBC336.06C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCB1_Taxid=4896;
RN NCB1_Taxid=4896;
RP SEQUENCE FROM N.A.
RC STRAIN=912;
RA Tozawa Y., Crouch R.J.;
RT "Genomic and cDNA sequences of Schizosaccharomyces pombe Ribonuclease H1."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Bozrym K., Beck A., Reinhardt R., McDougall R.C., Rajandream M.A.,
RA Barrell B.G.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS ENZYME IS AN ENDONUCLEASE THAT DEGRADATES THE RNA OF
CC -1- RNA-DNA HYBRIDS SPECIFICALLY.
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC -1- MONESTER.
CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RNASE H FAMILY.
DR EMBL; AF048992; AAC04366.1; -
DR EMBL; AL121815; CAB58158.1; -
DR PIR; T40244; T40244.
DR HSSP; Q04740; IQHK.
DR GenedB SPombe; SPBC336.06C; -
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:000287; F:magnesium ion binding; IEA.
DR GO; GO:0004524; F:ribonuclease H activity; IEA.
DR InterPro; IPR009027; L9_N like.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; RNaseH; 1.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding.
FT DOMAIN 55 62 POLY-SER.
FT VARIANT 78 82 POLY-SER.
FT VARIANT 129 129 D -> E (IN STRAIN 912).
FT METAL 129 129 MAGNESIUM (BY SIMILARITY).

```

```

FT METAL 171 171 MAGNESIUM (BY SIMILARITY).
FT METAL 191 191 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 264 AA; 29421 MW; 9838C8344B731DE0 CRC64;

Query Match 20.9%; Score 323.5; DB 3; Length 264;
Best Local Similarity 29.7%; Pred. No. 8,1e-22;
Matches 84; Conservative 41; Mismatches 105; Indels 53; Gaps 5;

QY 28 FYAVRGRKGTGFLVWNEGRQVDRFPAPRPFKATBEDANAFVR----- 72
DB 8 FYAVARGRGTGISTWDEASDVQKGYGNRYKKFSTYEAQDFCTEGSRYSSSGPPYR 67
QY 73 -----KSASPEVSEGHENQGE-SEAKPKRLREPLDGDGHSAGPVAKMKRS 121
DB 68 STTSYGYSPYSSSSNYSARISDKYRKISRSYTEKQIEIFSNDTHEKS----- 117
QY 122 VEPAPVSRDTFSPYNGDFVYVYTDGSSNSGRRKPRAGIGVYWGPGHPLNVGIRLPG-RQ 180
DB 118 -----IACSDRQVYVYADSSLRNGKKGAVAGCGVFFGDDPDPRNTISVPLAGEBQ 165
QY 181 TNGRAEIHACKAIEQAKTQINIKLVLTDSMFTINGITNNVQGWKKGWTKTSAGKEVIN 240
DB 166 TNNRAELQIILALNTS---GDLTIRSDSYSTKSILTWLPKKKNDFKTSNSQPVKN 221
QY 241 KEDFVALERLTQGMIDQMHHVPGHSGFIGNEADRLAREGAKQ 283
DB 222 LDILIRASDLMSDRVSVLEVVGHSTDYGNQADMLARGLASE 264

RESULT 9
Q8DM24 PRELIMINARY; PRT; 159 AA.
ID Q8DM24
AC 08DM24;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ribonuclease H.
GN RNHA OR TLR0299.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OC NCB1_Taxid=32046;
RN NCB1_Taxid=32046;
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005369; BAC07852.1; -
DR GO; GO:0003676; F:magnesium ion binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; RNaseH; 1.
KW Complete proteome.
SQ SEQUENCE 159 AA; 17790 MW; EE5900EC032931E2 CRC64;

Query Match 15.8%; Score 245; DB 16; Length 159;
Best Local Similarity 37.6%; Pred. No. 8,4e-15;
Matches 53; Conservative 26; Mismatches 52; Indels 10; Gaps 5;

QY 142 VYTDGSSNSGRRKPRP---GIGVYWGPGHPLNVGIRLPGRTNORPAEIHACKAIEQ-A 197
DB 7 IYTDGACBEGN---DQPGGMGVIVYFTDQSVHELGGHPR-TTNNRMELQAAIEALKWVR 61
QY 198 KTONINKVLVTDSMFTINGITNNVQGWKKGWTKTSAGKEVINKEDFVALERLTQGMIDQ 257
DB 62 QLAPGSAIALVYDSEVRLGITEWHHWRKGMKTAACKPVLNODUMOBELDALNDPL-VQ 120
QY 258 WNVHVPCHSGFIGNEADRLAR 278

```

Db 121 WHVRGHRGVDGNERCDLIAR 141

## RESULT 10

089U03

PRELIMINARY; PRT; 154 AA.

AC 089U03; 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Ribonuclease H.

GN BR1316.

Bradyrhizobium japonicum.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.

OX NCBI\_TaxID=375;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=USDA 110;

MEDLINE=22484998; Pubmed=12597275;

RX Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Matsumoto M., Idegawa K., Iriuchih M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tsubata S.;

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 Bradyrhizobium japonicum USDA110."

RL DNA Res. 9:189-197(2002).

DR EMBL; AP005939; BAC46581.1;

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0004523; F:ribonuclease H activity; IEA.

DR InterPro; IPR002156; RNaseH.

DR Pfam; PF00075; rnaaseH; 1.

DR Complete proteome.

SQ SEQUENCE 154 AA; 17264 MW; 580A32B67E57F414 CRC64;

Query Match 15.6%; Score 241.5; DB 16; Length 154;  
 Best Local Similarity 37.3%; Pred. No. 1.7e-14;  
 Matches 57; Conservative 18; Mismatches 49; Indels 29; Gaps 4;

QY 140 VVVYTGCCSSNGRRKPRAGIGVYWGPG-----HPLNVGIRLPGROTNRARI 187  
 Db 7 VVITDAGCGN-----GPGGALIKPKDKKELNGSR---HTTNOMEL 51  
 QY 188 HAACAIQAQKTONIKLVLYTDSMTFINGITNMVQGMKNGKTSAGKEVINKEDFVAL 247  
 Db 52 MAATSALEALKRP--CTVDLYTDSQVYRGITGMHGMKNKWRATADKKFVKVETLWQRL 109  
 QY 248 ERLTQGMIDQMHWVPGSGFTGNEADRLARREG 280  
 Db 110 DAALKAHQVRMHWKGAHGPENERADQLARDG 142

## RESULT 11

088UB4

PRELIMINARY; PRT; 298 AA.

AC 088UB4; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ribonuclease H (Putative) (EC 3.1.26.4).

GN RNH OR LP2593.

OS Lactobacillus plantarum.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
 OC Lactobacillus.

OX NCBI\_TaxID=1590;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=NCIMB 8826 / WCF51;  
 MEDLINE=22480296; Pubmed=12566566;  
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,  
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,  
 RA Fiers W.E.J., Strlekena W., Klein Lankhorst R.M., Bron P.A.,

RA Hofer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ureing B.,  
 RA De Vos W.M., Siezen R.J.;

RT "Complete genome sequence of Lactobacillus plantarum WCF51."

RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).

DR EMBL; AL935259; CAD64863.1;

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0004524; F:ribonuclease H activity; IEA.

DR InterPro; IPR009027; L9\_N like.

DR InterPro; IPR002156; RNaseH.

DR Pfam; PF00075; rnaaseH; 1.

DR Hydroxase; Complete proteome.

SQ SEQUENCE 298 AA; 33023 MW; DBB0A369663D6D2E CRC64;

Query Match 15.6%; Score 241; DB 16; Length 298;  
 Best Local Similarity 27.2%; Pred. No. 4.6e-14;  
 Matches 74; Conservative 40; Mismatches 92; Indels 66; Gaps 11;

QY 28 FYAVRRGRTGVLTNNECRAQYDRFPARFKFATEDEAMAFVRKSASPEVSEGHENOH 87  
 Db 5 YVAVRRGROPGRYRTPEOTQGVSPQAYKSFTEKXADPFMAKASP----- 54  
 QY 88 GQSEAKPKGRLEPLDGGHESAQYPAKMKPSVEPAPVSRDTSYMGDFVYVYTDGC 147  
 Db 55 -----TRPA-----HSK-SISNOTPVNA-----ATVYTDG 80  
 QY 148 CSNGRRKPRAGIGVYWGPGHPLNVGIRLPGR-----QTNRARHIAACAIQA 197  
 Db 81 SRRTGN---VAGQHVHDDKAAVAYNIEMPDQLVTSAGEMGATNNRMEIMAFRLAEQL 137  
 QY 198 KT--QVINKLVLYTDSMTFINGIT--NMVQGMKNGKTSAGKEVINKEDFVALERT--- 251  
 Db 138 QQLGQTKTGILFVLDSQVYLVNATKGLAGKRRGKRSNG--PLVNAELMREVDRLLPAF 196  
 QY 252 QGMIDQMHWVPGSGFTGNEADRLARREGAKQ 283  
 Db 197 TALNYR--TKGHATNGNVFVDHLNQMMDQ 225

## RESULT 12

0820F5

PRELIMINARY; PRT; 302 AA.

AC 0820F5; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ribonuclease H.

GN ALR0142.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

OX NCBI\_TaxID=103690;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=21595285; Pubmed=11759840;

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Matsumoto M., Iriuchih M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;

RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120."

RL DNA Res. 8:205-213(2001).

DR EMBL; AP003581; BAB77666.1;

DR PIR; AF1824; AF1824.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0004523; F:ribonuclease H activity; IEA.

DR InterPro; IPR002156; RNaseH.

DR Pfam; PF00075; rnaaseH; 1.

DR Complete proteome.

SQ SEQUENCE 302 AA; 33624 MW; F1C25DBE6E56D0752 CRC64;

Query Match 15.4%; Score 238; DB 16; Length 302;  
 Best Local Similarity 39.0%; Pred. No. 8.9e-14;

```

Matches      57;   Conservative      19;   Mismatches      50;   Indels      20;   Gaps      6
Qy          142 VYTGCCSSNRRKRPRA---GIYYWGFHPLNYGIRLPGROTNPAAIHAACAIE--- 195
Db          10 IYTDGACTGN----FGPGMGVGVVVFSDGSYHEMG--DAKHITNNKMEQAIAALAKFLH 64
Qy          196 ---DAKTQNINKVLVTDSMFTINGITNWGWKKNGKTSAGKEVINKEDEPVALERLTQ 252
Db          65 DSGGLE-----PILTYTSEYLINCVTVKWKGKKKKGGKSNDGPVNODLETIDEL-N 118
Qy          253 GMDIQMHVPGHSGFIGNEGADRLAR 278
Db          119 SRKYVMHHVRHGSHSGIENRCDIVAR 144

RESULT 13
Q7VM15      PRELIMINARY; PRT; 153 AA.
AC Q7VM15.
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Ribonuclease H.
GN RNHA OR HD1206.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730; [1]
RN COMPLETE FROM N.A.
RC STRAIN=3500HP / ATCC 700724;
RA Munson R.S., Jr., Ray W.C., Mahiriz G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
EMBL; AE017154; AAP96053.1; -.
KM Complete proteome.
SQ SEQUENCE 153 AA; 17453 MW; 390C8F87B24FD1D1 CRC64;

Query Match 14.3%; Score 221.5; DB 16; Length 153;
Best Local Similarity 35.1%; Pred. No. 1,2e-12;
Matches 54; Conservative 19; Mismatches 52; Indels 29; Gaps 4
Qy          140 VVVYTGGCCSSNGRRRKPAGIGVYWGPQHPLNVGIRLPGRQ-----TNQRAEI 187
Db          4 VNIFPDGCLGNP-----GPG---DIGVTLRVNQHQKKVSQGYPQTNNRMEL 48
Qy          188 HAACAIIQAKTQNIINKVLVTDSMFTINGITNWGWKKNGKTSAGKEVINKEDEPVALL 247
Db          49 RAIVLEGLMLK--EACNTVLTSDSQMYKNGITKWIPIFKKKSNWKYANGKAVXNKDLWILL 106
Qy          248 ERLTGMDIOWMHVPHSGFIGNEGADRLAREGA 281
Db          107 DEKIOIHTEMKVKVGHSGHYENEICDELAKLGA 140

RESULT 14
O82XV8      PRELIMINARY; PRT; 161 AA.
AC O82XV8.
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Probable ribonuclease h1 protein (EC 3.1.26.4).
GN RNHA OR NEO140.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915; [1]
RN COMPLETE FROM N.A.
RC STRAIN=ATCC 19718 /IFO 14298;
MEDLINE=22586410; PubMed=12700255;
EX
```

ID	Q081T47	PRELIMINARY;	PRT;	149 AA.
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Ribonuclease HI.			
OS	Leishmania major.			
CC	Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.			
NCBI	_taxid=5664;			
RV	[1]			
RA	SEQUENCE FROM N.A.			
RP	Misra S., Chaudhuri G.;			
RT	"Cloning and characterization of ribonuclease HI (RNase HI) from			
RT	Leishmania major.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF542055; AAN17340.1; "			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	GO; GO:0004523; F:ribonuclease H activity; IEA.			
DR	InterPro; IPR002156; RNaseH.			
PFam;	PF00075; RNaseH; 1.			
KW	Hydrolase; Complete proteome.			
SO	SEQUENCE 161 AA; 18195 MW; F7777C1947CFB49F CRC64;			
Query Match	14.3%; Score 221.5; DB 16; Length 161;			
Best Local Similarity	34.2%; Pred. No. 1.3e-12;			
Matches	53; Conservative 24; Mismatches 51; Indels 27; Gaps 5;			
QY	140 VVYTTDGCSSNGRRKRRPAGIGVYWGEGHPLNVGIRLGR-----QTNRATETHA 189			
DB	10 VEITTDGACKGN-----PGIG-GWG-----VCIKDGVEYREFGEPVTTNRMETLA 56			
QY	190 ACKRIEQAQAK-----TQNIINKLVLYTDSMFTINGITNWQWKGKMGTSAGKEVINKEDEFV 245			
DB	57 ATRALQALAESLPDTGSLRVQLHTSDQYVQKGISSEWHSMKRGMLTDKKPVAKEALWK 116			
QY	246 ALERLTQGMIDIOMNHVPGHSGFIGNEEADRLAREG 280			
DB	117 ELDDLSRRYQEWFWVVGHNCHDGNEDADMLANRG 151			
RESULT 15				
ID	Q081T47			
AC	Q081T47;			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Ribonuclease HI.			
OS	Leishmania major.			
CC	Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.			
NCBI	_taxid=5664;			
RV	[1]			
RA	SEQUENCE FROM N.A.			
RP	Misra S., Chaudhuri G.;			
RT	"Cloning and characterization of ribonuclease HI (RNase HI) from			
RT	Leishmania major.";			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF542055; AAN17340.1; "			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	GO; GO:0004523; F:ribonuclease H activity; IEA.			
DR	InterPro; IPR002156; RNaseH.			
PFam;	PF00075; RNaseH; 1.			
KW	Hydrolase; Complete proteome.			
SO	SEQUENCE 161 AA; 18195 MW; F7777C1947CFB49F CRC64;			
Query Match	14.2%; Score 219.5; DB 5; Length 149;			
Best Local Similarity	37.2%; Pred. No. 1.8e-12;			
Matches	54; Conservative 20; Mismatches 58; Indels 13; Gaps 5;			
QY	140 VVYTTDGCSSNGRRKRRPAGIGVYWGEGHPLNVGIR-LPGHQ--TNRATETHAACKAIE 195			
DB	4 VTISDGDACKGN-----PGPG---CMGAVLVANGHGVKELFGGEANTYNNRMELTVAIEALR 56			
QY	196 QAKTQNIINKLVLYTDSMFTINGITNWQWKGKMGTSAGKEVINKEDEFVLERLTQGM 255			
DB	57 ALKRP--CKQVWVWTDSDQYVQKGISSEWHSMKRGMAKTAKKPVAQADIMQALDAAAABHD 114			
QY	256 IQMNHVPGHSGFIGNEEADRLAREG 280			
DB	115 ISWHVVGHNCHDGNEDADMLANRG 139			



```

RESULT 16
Q88Y70 PRELIMINARY; PRT; 150 AA.
AC 088Y70;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Ribonuclease H1.
GN RNHA OR P3PT03712.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=323;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Barry K., Uteback T., Van Aken S., Feldblyum T., Gwin M.,
RA Dodson R., Deboy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beaman M., Halt D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE016869; AA057181.1; -.
DR TIGR: P3PT03712; -.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro: IPR002156; RNaseH.
DR Pfam: PF00075; rnaaseH; 1.
KW Complete proteome.
SQ
SEQUENCE 150 AA; 16782 MW; 4ED1FC436407C77C CRC64;

Query Match 14.2%; Score 219.5; DB 16; Length 150;
Best Local Similarity 33.1%; Pred. No. 1.8e-12;
Matches 53; Conservative 20; Mismatches 58; Indels 29; Gaps 5;

QY 136 MGPVVVYVYTGCCSSNGRRPRAGIGVYNGP-----CHPLVNGIRLPGRQTQ 183
DB 1 MSDSVEMFTDGACKGN---PGPG---GWSGALLVCKGVEKELWGGERN---TNN 45
QY 184 RAEIHAACKAIEQAKTONINKLVYTDSEFTINGITVWQGMKKGKTSAGKEVINKED 243
DB 46 RMLLTGAIRLEBEIKR--CEVLTVTDSQVYMKGITMMVMVWKKRGKTKAKEVYKADL 103
QY 244 FVALERTLQGMIDIQMHVPHSGFTIGNEADRLAREGAKQ 283
DB 104 WQLLDEQVSRHTVQWQVVRGHIIGHGNERADQLANRGVDE 143

RESULT 17
Q88FF5 PRELIMINARY; PRT; 148 AA.
AC 088FF5;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Ribonuclease H1.
GN RNHA OR PP4142.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=160468;
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beaman M., Deboy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapfel E., Scanlan D., Tran K., Moazzes A.,
RA Uteback T., Rizzo M., Lee K., Kosack D., Moesli D., Medlar H.,
RA Lauber J., Stjepanovic D., Hohlefeld J., Streletz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duecherneolt A., Tuenmler B.,

```

```

RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL: AE016789; AA069724.1; -.
DR TIGR: PP4142; -.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro: IPR002156; RNaseH.
DR Pfam: PF00075; rnaaseH; 1.
KW Complete proteome.
SQ
SEQUENCE 148 AA; 16913 MW; CB869BAD513B8953 CRC64;

Query Match 14.1%; Score 218.5; DB 16; Length 148;
Best Local Similarity 35.5%; Pred. No. 2.2e-12;
Matches 54; Conservative 21; Mismatches 64; Indels 13; Gaps 4;

QY 136 MGPVVVYVYTGCCSSNGRRPRAG-----IGVWPGPHPLVNGIRLGRGTQRAEIHAC 191
DB 1 MSDSVEMFTDGACKGN---PGPGGVLWYIKGVEKELWGGRR---ETTNNMEIMAAI 53
QY 192 KATIEQAKTONINKLVYTDSEFTINGITVWQGMKKGKTSAGKEVINKEDFVALERTL 251
DB 54 QGLMSLRK--CEVLTVTDSQVYMKGINEMVMVWKKRGKTKAKEVYKADLMQQLDEOV 111
QY 252 QGMIDIQMHVPHSGFTIGNEADRLAREGAKQ 283
DB 112 NRKATVWQVVRGHIIGHGNERADQLANRGVDE 143

RESULT 18
Q934V8 PRELIMINARY; PRT; 202 AA.
AC Q934V8;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Putative ribonuclease H.
GN HCM2.0109C.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
NCBI_TaxID=601;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Kung'u K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852 (2001).
DR EMBL: AL513384; CAD09976.1; -.
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro: IPR002156; RNaseH.
DR Pfam: PF00075; rnaaseH; 1.
KW plasmid; Complete proteome.
SQ
SEQUENCE 202 AA; 23149 MW; CAF9E0A155E8ED9 CRC64;

Query Match 13.8%; Score 214; DB 16; Length 202;
Best Local Similarity 31.1%; Pred. No. 8.9e-12;
Matches 59; Conservative 26; Mismatches 75; Indels 30; Gaps 4;

QY 109 ESAQPYAKMKKSVVEPAPVSRDTSYMGDFVYVYTDGCCSSNGRRKPRAGIGVYNGPGH 168

```



```

DB      16 KSKARVYHQGSRNFKPTPVV--FNPOKTVKVSDDSCCLKN-----PGG 59
OY      169 PLVNGVIRLPGR-----QTNORAEIHAACKAIEQATONINKLVLTDSMFTTN 216
DB      60 PGYGVIVLQYRGEERELSGFSTTNNMEMGALIALERLKP--CNVILHSDSQYKLN 117
OY      217 GTTNVQCKKNGKMTSAGKEVINKEDEPVALERLQGMQMDIQMNYRGHSGFIGNEARL 276
DB      118 GTTLMMKGMKRNKMTSEKKPVKNVLDLPAASRHNVRKWKVGHAGHRENEKCDRL 177
OY      277 AREGAKOSED 286
DB      178 AKIAVSAAD 187

RESULT 19
OY      07WCJ8      PRELIMINARY;      PRT;      155 AA.
AC      07WCJ8;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE      Ribonuclease HI (EC 3.1.26.4).
GN      DASP OR HERA OR RNH OR RNHA OR SDRS OR BB4278.
OS      Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Alcaligenaceae; Bordetella.
OY      NCBI_TaxID=518;
RN      NCBI_TaxID=518;
RP      SEQUENCE FROM N.A.
RC      STRAIN=RB50 / ATCC BAA-588;
RX      MEDLINE=22827954; PubMed=12910271;
RA      Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA      Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA      Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA      Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA      Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA      Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA      Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA      Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA      Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA      Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT      "Comparative analysis of the genome sequences of Bordetella pertussis,
RT      Bordetella parapertussis and Bordetella bronchiseptica.";
RL      Nat. Genet. 35:32-40(2003).
DR      EMBL: BX640450; CAB3642.1; -.
KW      Hydrolyase; Complete proteome.
SQ      SEQUENCE 155 AA; 17455 MW; DF2709B4B8343CEF CRC64;

Query Match      13.6%; Score 209.5; DB 16; Length 155;
Best Local Similarity 33.6%; Pred. No. 1.6e-11;
Matches 50; Conservative 25; Mismatches 63; Indels 11; Gaps 4;

OY      137 GDFVVVYTDGSSNGRRKPRAGIGVYMGPGH---PLVNGVIRLPGRQTNORAEIHAACKA 193
DB      10 GQGVEMWTDGACKN---PGPGWGVLMRACQHEKTMHGER---QTNRMELMAVIBG 63
OY      194 IEQATONINKLVLTDSMFTTINGITTNVQCKKNGKMTSAGKEVINKEDEPVALERLTOG 253
DB      64 LRALKRP--CRVITHTDQYVWKMTETLANKRGMTRADKKPVKNVLMQALDBOYGR 121
OY      254 MDIQMNYRGHSGFIGNEARLAREGAK 282
DB      122 HQVQWRWVRGHAGDGNERRADALANQVGE 150

RESULT 20
OY      07WOT2      PRELIMINARY;      PRT;      155 AA.
AC      07WOT2;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)

```

```

DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Ribonuclease HI (EC 3.1.26.4).
GN      DASP OR HERA OR RNH OR RNHA OR SDRS OR BP2834.
OS      Bordetella parapertussis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Alcaligenaceae; Bordetella.
OY      NCBI_TaxID=519;
RN      NCBI_TaxID=519;
RP      SEQUENCE FROM N.A.
RC      STRAIN=12822 / ATCC BAA-587;
RX      MEDLINE=22827954; PubMed=12910271;
RA      Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA      Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA      Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA      Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA      Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA      Feltham T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA      Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA      Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA      Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA      Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT      "Comparative analysis of the genome sequences of Bordetella pertussis,
RT      Bordetella parapertussis and Bordetella bronchiseptica.";
RL      Nat. Genet. 35:32-40(2003).
DR      EMBL: BX640434; CAB39117.1; -.
KW      Hydrolyase; Complete proteome.
SQ      SEQUENCE 155 AA; 17455 MW; DF2709B4B8343CEF CRC64;

Query Match      13.6%; Score 209.5; DB 16; Length 155;
Best Local Similarity 33.6%; Pred. No. 1.6e-11;
Matches 50; Conservative 25; Mismatches 63; Indels 11; Gaps 4;

OY      137 GDFVVVYTDGSSNGRRKPRAGIGVYMGPGH---PLVNGVIRLPGRQTNORAEIHAACKA 193
DB      10 GQGVEMWTDGACKN---PGPGWGVLMRACQHEKTMHGER---QTNRMELMAVIBG 63
OY      194 IEQATONINKLVLTDSMFTTINGITTNVQCKKNGKMTSAGKEVINKEDEPVALERLTOG 253
DB      64 LRALKRP--CRVITHTDQYVWKMTETLANKRGMTRADKKPVKNVLMQALDBOYGR 121
OY      254 MDIQMNYRGHSGFIGNEARLAREGAK 282
DB      122 HQVQWRWVRGHAGDGNERRADALANQVGE 150

RESULT 21
OY      07VOB6      PRELIMINARY;      PRT;      167 AA.
AC      07VOB6;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE      Ribonuclease H (EC 3.1.26.4).
GN      RNHA OR BFL224.
OS      Candidatus Blochmannia floridanus.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OY      NCBI_TaxID=203907;
RN      NCBI_TaxID=203907;
RP      SEQUENCE FROM N.A.
RC      MEDLINE=22784745; PubMed=12886019;
RX      Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,
RA      Latorre A., Rauebel C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA      van Ham R.C.H.J., Gross R., Moya A.;
RT      "The genome sequence of Blochmannia floridanus: comparative analysis
RT      of reduced genomes.";
RL      Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).
DR      EMBL: BX248584; CAD83738.1; -.
KW      Hydrolyase; Complete proteome.
SQ      SEQUENCE 167 AA; 19480 MW; 3071B72795CC30ED CRC64;

Query Match      13.5%; Score 209; DB 16; Length 167;
Best Local Similarity 36.1%; Pred. No. 2e-11;

```

Matches	52; Conservative	25; Mismatches	61; Indels	6; Gaps
Qy	140 VVVTIDGCCSSNGRRKPRAGICVYVGPCHPLNVGIRLGRQTNQAAEIHAAKIAEOKAT			199
Db	5 IEIFDSDCLNPGPGGCAGLIRYQYKKESAGYHI--TTNNMELMAAIIALESK-			60
Qy	200 QINIKLVYVYDTSMEFTINGITWVQGMKKNQWTSAGKEVINKEDEVALERLTQGMIDQ-W			258
Db	61 -NSCGIILYDSQVYLITITQIQWKHKHMTATSKLVKNIDLMRRDLIAIQPINIKDW			119
Qy	259 MHVPGHSGFIGNEEADRLAREGAK	282		
Db	120 RWLKSHTGHPDNERCDQARAK	143		

  

RESULT 22	
ID Q7V474	PRELIMINARY; PRT; 257 AA.
AC Q7V474	
DT 01-OCT-2003 (TREMBLrel. 25, Created)	
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)	
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DR Possible ribonuclease HI (EC 3.1.26.4).	
GN RNAA OR PWT2093.	
OS Prochlorococcus marinus (Strain MIT 9313).	
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;	
OC Prochlorococcus.	
OX NCBI_TaxID=74547;	
RA RA	
RA MEDLINE=22825698; PubMed=12917642;	
RA Rocap G., Latimer F.W., Lamerdin J., Malfatti S., Chain P.,	
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,	
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,	
RA Shaw S.L., Seeglich C., Sullivan M.B., Ting C.S., Tolonen A.,	
RA Webb E.A., Zinser E.R., Chisholm S.W.;	
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic	
RT niche differentiation."	
RL Nature 424:1042-1047(2003).	
DR EMBL; BX572101; CAE22267.1; -	
KM Hydrolyse; Complete proteome.	
SEQ SEQUENCE 257 AA; 28057 MW; BD726C66801AA85 CRC64;	

  

Query Match	
Beet local similarity	37.3%; Score 209; DB 16; Length 257;
Matches 57; Conservative	20; Mismatches 50; Indels 26; Gaps 8;
Qy	137 GDFVVVVYDGCSSNGRRKPRAGICVYVNG-----PGHPLNVGILPGQTNQAAEIHAA 190
Db	6 GRVVAATAATDGAACSGN----PGGCG--GNGALLRFEDSGVEERGVGAIPA-TTNNRHELDAA 57
Qy	191 CKAIEQATQTNIN-KLVLYTDSMEFTINGITWVQGMKKNQWTSAGKEVINKEDEVALER 249
Db	58 LHVLEQLLELPCHPDLKIRTSKYILIDISKMAKMKRKGRTAAGKEVVLNODLMRALDR 117
Qy	250 LTQGMIDQWMAVP-----GHSGFIGNEEADRLA 277
Db	118 AR--LD---HVPPLAVYVKGHSGDPNERKVDQIA 144

  

RESULT 23	
ID Q7VRX8	PRELIMINARY; PRT; 155 AA.
AC Q7VRX8	
DT 01-OCT-2003 (TREMBLrel. 25, Created)	
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)	
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DE Ribonuclease HI (EC 3.1.26.4).	
GN DASF OR HERA OR RNH OR RNHA OR SDBA OR BP3211.	
OS Bordetella pertussis.	
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;	
OC Alcaligenaceae; Bordetella.	
OX NCBI_TaxID=520;	

[1]

```
RN RESULT 24
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parikh H.J., Seibidha M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerceno-Terriza A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bacon N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertcak H., O'Neill S., Ormond D., Price C.,
RA Raab-Hovav E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Bordetella parapertussis and Bordetella bronchiseptica.";
RD Nat. Genet. 35:32-40(2003).
DR EMBL; BX640420; CAE3477.1; -.
RW Hydrolyase; Complete proteome.
SQ SEQUENCE 155 AA; 17487 MW; DF2709BA8B3428BF CRC64;

Query Match 13.5%; Score 208.5; DB 16; Length 155;
Best Local Similarity 34.0%; Pred. No. 2e-11; Matches 11; Gaps 4
Matches 50; Conservative 24; Mismatches 62; Indels 11; Gaps 4

OY 137 GDPVVVYTDGCCSSNGRRKPRAGIGVYGPH--PLNVGIRLPGRQTNORAEIHACKA 193
DB 10 GOVEWMTDGACKGN---PGSGGVCLMRAGQHKEKTHMGERR--QTNNRMELMAVIEG 63
OY 194 IEQAQTQINIKLVLYTDSMFTINGITNNVOGKKRGTSTSAKGEIYNEDFVALERLTQG 253
DB 64 LRALKRP--CVTTHTDTSQVMKKGTELANWKRGKGMETADKPKNVELMQALDEQVGR 121
OY 254 MDIQMHVPYHSGFGFIGNEADRLAREG 280
DB 122 HQVQWRWVRGHAGDPGNERDALANNQG 148

AC Q7U3U1 PRELIMINARY; PRT; 248 AA.
ID Q7U3U1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Possible ribonuclease HI (EC 3.1.26.4).
GN RNHA OR SYNW2338.
OS Synecococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chlorococcales; Synecococcus.
CC NCBI_TaxID=84589;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahamsia B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarron J., Paulsen I.,
RA Dufresne A., Patensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569695; CAB08853.1; -.
RW Hydrolyase; Signal; Complete proteome.
PT SIGNAL 1 Potential.
SQ SEQUENCE 248 AA; 27400 MW; BDF4CA6106FA703 CRC64;

Query Match 13.5%; Score 208; DB 16; Length 248;
Best Local Similarity 36.7%; Pred. No. 4.3e-11;
Matches 55; Conservative 23; Mismatches 52; Indels 20; Gaps 7

OY 137 GDPVVVYTDGCCSSNGRRKPRAGIGVYG-----PGHPVLNVGIRLPGRQTNORAEIHAA 190
DB 6 GRVVAARADGACSGN----PGRG---GMGALLRFEDSDVEERFGCHDPA-TTNNRMELQAA 57
OY 191 CKAIHQAK-TQNIINKLVLYTDSMFTINGITNNVOGKKRGTSTSAKGEIYNEDFVALERLTQA 248
```

Db 58 LELLQRLKQLPRHPDLTIRTDKSKYLIDGLGSMWKGKRGKTAAGKPYLNQDLWKALDA 117  
QY 249 -RLTQGMIDIQMHVPGHSGFIGNEDRLA 277  
Db 118 ARLD---DVPLAYVKGHSGDPDNDVDRYA 144

## RESULT 25

Q8EE30 PRELIMINARY; PRT; 158 AA.  
AC Q8EE30;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Ribonuclease HI.  
GN RNHA OR S02560.  
OS Shewanella oneidensis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Alteromonadaceae; Shewanella.  
OX NCBI\_TaxId=70863;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MR-1;  
RX MEDLINE=22297686; PubMed=12368813;  
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
Read T.D., Eisen J.A., Seshadri R., Ward N., Mehe B., Clayton R.A.,  
Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,  
Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
Vamathevan J., Weidman J., Impriali M., Lee K., Berry K., Lee C.,  
Mueller J., Khouli H., Gill J., Uterback T.R., McDonald L.A.,  
RA Feldblyum T.V., Smith H.O., Venter J.C., Nelson K.H., Fraser C.M.,  
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
Shewanella oneidensis.";  
RT Nat. Biotechnol. 20:1118-1123(2002).  
RL EMBL; AB015697; AAN55590.1; -.  
DR TIGR; S02560; -.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR InterPro; IPR002156; RNaseH.  
DR Pfam; PF00075; rnaseh; 1.  
KW Complete proteome.  
SQ SEQUENCE 158 AA; 17782 MW; BDEBAE51CC162E4B CRC64;

Query Match 13.2%; Score 203.5; DB 16; Length 158;

Best Local Similarity 35.4%; Pred.No.6.1e-11;

Matches 51; Conservative 19; Mismatches 63; Indels 11; Gaps 4;

QY 142 VTTDCCSSNGRRKPRAGIGV--YWGPHPLNVGIRLPGRTNORAEIHAACKAIEQAK 198  
Db 9 ITDSSCLGN---PGGGYGIWVNYKGTKEKSDGFSL---TTNNMEILAPIVALALK 62  
QY 199 TONINKLVLYTDSMPTINGITWVQGMKQKWTSAKEVINKEDFVALERLTQGMIDQW 258  
Db 63 BP--CKIILTSDSQVWRGIMTWIHGKKKGMWTSNRTPTVKNVDLWKRDKAAQLHQIDW 120  
QY 259 MNPVGHSGFTGNEDRLAREGAK 282  
Db 121 RWVKGHAGHAENERCQDLARAAAE 144

Search completed: October 7, 2004, 08:07:52  
JOB time : 127 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:02:55 : Search time 24 seconds

(without alignment) 620.503 Million cell updates/sec

Title: US-10-054-313-1

Perfect score: 1546

Sequence: 1 MSWLLFLAHRYVALALAPCR.....FIGNEADRLAREGAKXSD 286

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : SWISSPROT42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1526	98.7	286	1	RNH1_HUMAN
2	1197.5	77.5	285	1	RNH1_MOUSE
3	299.5	19.4	494	1	RNH1_CRIFA
4	249.5	16.1	348	1	RNH1_YEAST
5	240.5	15.6	154	1	RNH_PASMU
6	239	15.5	169	1	RNH_TREPA
7	238.5	15.4	154	1	RNH_BRUME
8	236.5	15.3	154	1	RNH_HABIN
9	236.5	15.3	177	1	RNH_RHILLO
10	233.5	15.1	148	1	RNH_PSEAE
11	232.5	15.0	145	1	RNH_NEIMA
12	229.5	14.8	145	1	RNH_NEIMB
13	224.5	14.5	156	1	RNH_ZYMO
14	222.5	14.4	153	1	RNH_RHIME
15	221.5	14.3	153	1	RNH_HAEDU
16	220.5	14.3	146	1	RNH_AGRTS
17	219.5	14.2	149	1	RNH_CAUCR
18	219.5	14.2	150	1	RNH_PESBM
19	218.5	14.1	148	1	RNH_PSEPK
20	217.5	14.1	150	1	RNH_XANCP
21	216.5	14.0	151	1	RNH_RALSO
22	216	14.0	159	1	RNH_MYCSM
23	214.5	13.9	154	1	RNH_YERPE
24	214	13.8	156	1	RNH_VIBCH
25	209.5	13.6	155	1	RNH_VIBVU
26	208.5	13.5	155	1	RNH_ECOLI
27	208.5	13.5	156	1	RNH_PHOLU
28	206.5	13.4	154	1	RNH_VIBPA
29	206.5	13.4	155	1	RNH_SALTY
30	203.5	13.2	158	1	RNH_SHEON
31	199.5	12.9	150	1	RNH_XANAC
32	198.5	12.8	157	1	RNH_WIGBR
33	195.5	12.6	150	1	RNH_XYLFT

34	194.5	12.6	150	1	RNH_XYLFA	O9pb16 xyliella fas
35	189	12.2	231	1	RNH_STRCO	O9x7c6 streptomyce
36	188.5	12.2	152	1	RNH_RICPR	O9x7c3 rickettsia
37	183.5	11.9	152	1	RNH_RICCN	O9x915 rickettsia
38	179	11.6	153	1	RNH_THETN	O8ra67 thermocyc
39	174.5	11.3	160	1	RNH_SYNY3	O55801 synectocyst
40	173.5	11.2	153	1	RNH_BUCBP	P59434 buchnera ap
41	171.5	11.1	161	1	RNH_BUCAP	O08885 buchnera ap
42	171	11.1	156	1	RNH_THETH	P29253 thermus the
43	163	10.5	1189	1	POL_BAEVM	P10272 baboon endo
44	160	10.3	1165	1	POL_GALV	P21414 gibbon ape
45	155	10.0	1204	1	POL_MLVFP	P26808 firend murti
46	152	9.8	1204	1	POL_MLVFP	P26809 firend murti
47	150	9.7	1199	1	POL_MLVMO	P03355 moloney mur
48	147	9.5	1046	1	POL_PENVA	P31792 feline endo
49	147	9.5	1204	1	POL_MLVFS	P26810 firend murti
50	145	9.4	1196	1	POL_MLVRD	P11227 radiation m
51	143	9.2	843	1	POL_MLVAV	P03357 akr murine
52	143	9.2	1196	1	POL_MLVAV	P03356 akr murine
53	142	9.2	146	1	RNH_CAMTE	O9pm39 campylobact
54	141.5	9.2	1049	1	POL_FIVPE	P18042 human immun
55	140.5	9.1	1124	1	POL_FIVPE	P16088 feline immu
56	139.5	9.0	1124	1	POL_FIVSD	P15028 feline immu
57	138.5	9.0	1073	1	POL_HV2D1	P17757 human immun
58	138	8.9	581	1	POL_MLVK	P31795 radiation m
59	135.5	8.8	1124	1	POL_FIVTE	P31822 feline immu
60	134.5	8.7	1055	1	POL_HV2ST	P20876 human immun
61	134	8.7	1101	1	POL_VILVK	P35956 viena lenti
62	133	8.6	1086	1	POL_OMVVS	P16901 ovine lenti
63	133	8.6	1105	1	POL_VILV	P03370 viena lenti
64	133	8.6	1105	1	POL_VILV2	P23427 viena lenti
65	132.5	8.6	1034	1	POL_HV2CA	P24107 human immun
66	132.5	8.6	1035	1	POL_HV2SB	P12451 human immun
67	129.5	8.4	1056	1	POL_BIV06	P19560 bovine immu
68	129.5	8.4	1056	1	POL_BIV27	P19561 bovine immu
69	129	8.3	1105	1	POL_VILV1	P23426 viena lenti
70	128.5	8.3	1142	1	POL_VILV1	P18096 human immun
71	127.5	8.2	1035	1	POL_HV2KR	O74120 human immun
72	127.5	8.2	1035	1	POL_HV2NZ	P05962 human immun
73	125.5	8.1	1036	1	POL_HV2RO	P04584 human immun
74	125.5	8.1	1056	1	POL_SIVM1	P05896 simian immu
75	122	7.9	1047	1	POL_SIVM1	P27973 simian immu
76	120.5	7.8	1046	1	POL_SIVAG	P27980 simian immu
77	119.5	7.7	867	1	POL_IPMA	P13368 mouse intra
78	117.5	7.6	1161	1	POL_SIVL	P23074 simian foam
79	117	7.6	1109	1	POL_CAESC	P33459 caprine art
80	116.5	7.5	1061	1	POL_SIVAT	P05895 simian immu
81	115.5	7.5	886	1	POL_POAMV	P14350 human spuna
82	114.5	7.4	1027	1	POL_SIVCZ	P12783 chimpanzee
83	112.5	7.3	522	1	IBMF_CAMVB	P16666 caulliflowe
84	112.5	7.3	1058	1	POL_HV2D2	P15833 human immun
85	111.5	7.2	143	1	RNH_HELPY	P56120 helicobacte
86	109.5	7.1	1019	1	POL_SIVS4	P14502 simian immu
87	108.5	7.0	1022	1	POL_SIVSP	P19505 simian immu
88	108	7.0	1157	1	POL_SEVJL	P27401 simian foam
89	107.5	7.0	1057	1	POL_SIVAT	O02836 simian immu
90	107	6.9	1002	1	POL_HV1MA	P04588 human immun
91	107	6.9	1002	1	POL_HV1U4	P24740 human immun
92	105.5	6.8	1145	1	POL_EIAYV	P033740 human immu
93	105	6.8	1015	1	POL_HV1PV	P03368 human immu
94	104.5	6.8	143	1	RNH_HELPD	O9x1h3 helicobacte
95	104.5	6.8	1146	1	POL_EIAYV	P13204 equine infe
96	104.5	6.8	1015	1	POL_EIAYV	P32542 equine infe
97	102	6.6	1015	1	POL_HV1BS	P04587 human immu
98	102	6.6	814	1	POL_HV1BR	P03367 human immu
99	101	6.5	814	1	POL_IPMA1	P18894 mouse intra
100	101	6.5	1015	1	POL_HV1B1	P03366 human immu

## ALIGNMENTS

RESULT 1

RNH1\_HUMAN STANDARD; PRT; 286 AA.  
 ID RNH1\_HUMAN 060930; 060523; 060857;  
 AC 060930; 060523; 060857;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ribonuclease H1 (EC 3.1.26.4) (RNase H1) (Ribonuclease H type II).  
 GN RNASEH1 OR RNH1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RX MEDLINE=99017966; PubMed=9799596;  
 RA Cerritelli S.M., Crouch R.J.;  
 RT "Cloning, expression, and mapping of ribonucleases H of human and  
 mouse related to bacterial RNase H1.";  
 RL Genomics 53:300-307(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98171047; PubMed=9512096;  
 RA Wu H., Lima W.F., Crooke S.T.;  
 RT "Molecular cloning and expression of cDNA for human RNase H.";  
 RL Anticancer Nucleic Acid Drug Dev. 8:53-61(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Frank P., Braunschweig-Reiter C., Wintersberger U.;  
 RT "Cloning and functional expression of human RNase H1.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klotzner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Halel P.,  
 RA Diatchenko L., Marusik A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,  
 RA Bosak S.A., McEwen K.J., McKernan K.J., Malek J.A., Gamarallane P.H.,  
 RA Rickards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hult S.W.,  
 RA Vallilou D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Feihly J., Hulton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skaleka U., Small D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP CHARACTERIZATION.  
 RX MEDLINE=99428493; PubMed=10497183;  
 RA Wu H., Lima W.F., Crooke S.T.;  
 RT "Properties of cloned and expressed human RNase H1.";  
 RL J. Biol. Chem. 274:28270-28278(1999).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrid specifically.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonocester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit.  
 CC -1- ENZYME REGULATION: In the presence of magnesium, manganese is  
 CC inhibitory.  
 CC -1- SUBUNIT: Monomer (probable).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL: AF048995; AAC78564.1; -;  
 DR EMBL: AF048994; AAC78563.1; -;  
 DR EMBL: AF039652; AAC09261.1; -;  
 DR EMBL: AJ224117; CA11835.1; -;  
 DR EMBL: BC002873; AA002873.1; -;  
 DR HSP; Q04740; IQH.  
 DR Genew; HGNC:18466; RNASEH1.  
 DR MIM; 604123; -;  
 DR GO; GO:0004540; F:ribonuclease activity; TAS.  
 DR GO; GO:0006401; P:RNA catabolism; TAS.  
 DR InterPro; IPR009027; I9\_N-like.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnaeh; 1.  
 KM Hydrolyase; Nuclease; Endonuclease; Magnesium.  
 FT METAL 145 145 MAGNESIUM (BY SIMILARITY).  
 FT METAL 186 186 MAGNESIUM (BY SIMILARITY).  
 FT METAL 210 210 MAGNESIUM (BY SIMILARITY).  
 FT METAL 274 274 MAGNESIUM (BY SIMILARITY).  
 FT CONFLICT 4 4 L -> F (IN REF. 1).  
 FT CONFLICT 89 89 Q -> R (IN REF. 3).  
 FT CONFLICT 223 223 Q -> R (IN REF. 3).  
 SQ SEQUENCE 286 AA; 32064 MW; 400FE04E7B5CA6A CRC64;  
 Query Match 98.7%; Score 1526; DB 1; Length 286;  
 Best Local Similarity 98.6%; Pred. No. 6,6e-128;  
 Matches 282; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MSWLFPAHVAALPCRRGSGFGMFAVVRGRGTGVLTTNEPRAQVDRPAPRFX 60  
 DB 1 MSWLFPAHVAALPCRRGSGFGMFAVVRGRGTGVLTTNEPRAQVDRPAPRFX 60  
 QY 61 FATEDENAFVRKASPEVSEGHENHGOSEAKPKRLREPLDGDHESAGQPYAKMKP 120  
 DB 61 FATEDENAFVRKASPEVSEGHENHGOSEAKPKRLREPLDGDHESAGQPYAKMKP 120  
 QY 121 SVBPAPVSDTSTYWGDFPVVYTTDCCSNGRRKPRAGIGVYWGCHPLNIGRLPRQ 180  
 DB 121 SVBPAPVSDTSTYWGDFPVVYTTDCCSNGRRKPRAGIGVYWGCHPLNIGRLPRQ 180  
 QY 181 TNGRAETHAACKIAEOKATONIKLVLYTDSMTTGTITWVGWKKNGKTSAGKEVIN 240  
 DB 181 TNGRAETHAACKIAEOKATONIKLVLYTDSMTTGTITWVGWKKNGKTSAGKEVIN 240  
 QY 241 KEDFVALERLTQGMIDQMHVPCHSGFIGNEBADRLAREGAKOSED 286  
 DB 241 KEDFVALERLTQGMIDQMHVPCHSGFIGNEBADRLAREGAKOSED 286  
 RESULT 2  
 RNH1\_MOUSE STANDARD; PRT; 285 AA.  
 ID RNH1\_MOUSE 070338;  
 AC 070338;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribonuclease H1 (EC 3.1.26.4) (RNase H1).  
 GN RNASEH1 OR RNH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RX MEDLINE=99017966; PubMed=9799596;  
 RA Cerritelli S.M., Crouch R.J.;

RT "Cloning, expression, and mapping of ribonucleases H of human and mouse related to bacterial RNase H.",  
 RT Genomics 53:300-307(1998).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
 CC -1- ENZYME REGULATION: In the presence of magnesium, manganese is  
 CC inhibitory (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF048993; AAC78562.1; -.  
 CC DR HSSP: Q04740; 1QHK.  
 CC DR MGD: MGI:1335073; RNaseH1.  
 CC DR InterPro: IPR009027; L9\_N\_1like.  
 CC DR InterPro: IPR002156; RNaseH.  
 CC DR Pfam: PF00075; RNaseH; 1.  
 CC KM Hydroxylase; Nuclease; Endonuclease; Magnesium.  
 CC FT METAL 144 144 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 185 185 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 209 209 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 273 273 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 285 AA; 31805 MW; 92B09F8EFA822F CRC64;  
 Query Match 77.5%; Score 1197.5; DB 1; Length 285;  
 Best Local Similarity 77.3%; Pred. No. 8.4e-99;  
 Matches 221; Conservative 27; Mismatches 37; Indels 1; Gaps 1;  
 QY 1 MSWLLFLAHRVALALPCRSGRGFMFYAVRRGRKTGVFLTNWECRAQVDRPPARPKK 60  
 DB 1 KMWLLPLSTVTLAVRLRGICGLMFYAVRRGRGTGVFLSWSECKAQVDRFPARPKK 60  
 QY 61 FATEDEAMAFVRKSAPEVSEBGENHOGESAKPEKRLRPLDGGHSDAPYAKMMP 120  
 DB 61 FATEDEAMAFVRSSSPDSSKQESAEKQSAKTSKRPREPL-GESELPPEGPRTKQ 119  
 QY 121 SVEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRKPRAGIGVYWGPHPLANGIRLPGRO 180  
 DB 120 DTEPAPVSVKOTFSTMGESVITYTDGCCSSNGRRKARAGIGVYWGPHPLANGIRLPGRO 179  
 QY 181 TNQRAEIHAAKCAIEQAKTONINKLVLYTDSMTFINGITNMYQGMKNGKWTSAKEVIN 240  
 DB 180 TNQRAEIHAAKCAIMQAKQONISKLVLTYDSMTFINGITNMYQGMKNGKWTSTGKDVIN 239  
 QY 241 KEDFPALERTLQGMIDQMGHVHSGFGFINGEEDRLARAGAKOSD 286  
 DB 240 KEDFELDELQGMIDQMGHIVPHSGFVGNEDRLARAGAKOSD 285  
 RESULT 3  
 RNH1\_CRIFA STANDARD; PRT; 494 AA.  
 AC 007762;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribonuclease H (EC 3.1.26.4) (RNase H).  
 GN RNH1.  
 OS Crithidia fasciculata.  
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.  
 OX NCBI\_TaxId=5656;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C1;  
 RX MEDLINE=94022373; PubMed=8415705;  
 RA Campbell A.G., Kay D.S.;  
 RT "Functional complementation of an Escherichia coli ribonuclease H  
 RT mutation by a cloned genomic fragment from the trypanosomatid  
 RT Crithidia fasciculata";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:9350-9354(1993).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L18916; AAA03546.1; -.  
 CC DR PIR: A48683; A48683.  
 CC DR HSSP: Q04740; 1QHK.  
 CC DR InterPro: IPR009027; L9\_N\_1like.  
 CC DR InterPro: IPR002156; RNaseH.  
 CC DR Pfam: PF00075; RNaseH; 1.  
 CC KM Hydroxylase; Nuclease; Endonuclease; Magnesium.  
 CC FT METAL 281 281 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 325 325 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 374 374 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 494 AA; 53704 MW; 3E097DD1F76BE78 CRC64;  
 Query Match 19.4%; Score 299.5; DB 1; Length 494;  
 Best Local Similarity 27.4%; Pred. No. 6.2e-19;  
 Matches 98; Conservative 44; Mismatches 121; Indels 95; Gaps 10;  
 QY 10 RVALLALPCRSGRGFMFYAVRRGRKTGVFLTNWECRAQVDRPPARPKKATEDEAMA 69  
 DB 141 RTSCAPPP---ASRMKPSFYVAVVAVRGRIYSTWDQCEQVFGAVYKSPRTSEARA 197  
 QY 70 FVRKASPEVSEBGENHOGESAKPEKRLRPLDGGHSDAPYAKM 118  
 DB 198 YL--TAHPARSGLEKSDGDAASLSALSEPVGRLRSRAAEASVYVEADAPQLROR 255  
 QY 119 KESVEP--APVSRDTFSYMGDFVYVYTDGCCSSNGRRKPRAGIGVYWG--PGHPLANG 173  
 DB 256 VEEVPSGAAAVQRAESSVP---QVYVYDGCASHNGTRKARAGYGGFYGSTSDSNFSLP 312  
 QY 174 IRLPGRTNORAEIHAAKCAIEQA-----KTQINKLVLY 208  
 DB 313 VPTEAQTNNGEMKAVIHICIQGFVDAGVPPALGTSHCVPEPWELSELPPRLRLVLY 372  
 QY 209 TDSMTFINGITNMYQGMKNGKWTSAKEVINKEPVALERL----- 250  
 DB 373 TDSRYVIDELTRYALKVWANGFKLASKEPVLNQDMRLQIRLRDAYNTRYAQQHMAAT 432  
 QY 251 -----TOGMIDQMGHVHSGFGFINGEEDRLARAGAK 282  
 DB 433 GSHASTRPAPASQSRPFTHNTRNDETGEIELR--HYGKSHNDYGNENADVLAVAGAR 488  
 RESULT 4  
 RNH1\_YEAST STANDARD; PRT; 348 AA.  
 AC 004740;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease H (EC 3.1.26.4) (RNase H).  
 GN RNH1\_YEAST  
 OS Saccharomyces cerevisiae.  
 OC Eukaryota; Euzoenozoa; Fungi; Basidiomycota; Ascomycota; Saccharomycetes; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4932;  
 RN [1]

GN RNH1 OR YMR234W OR YMR959.16.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288c / AB972;  
 RX MEDLINE=9733268; PubMed=9169872;  
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
 RA Jagsel K., Iye G., Moule S., Odell C., Pearson D., Rajadream M.A.,  
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 XII.";  
 RL Nature 387:90-93(1997).  
 RN [2]  
 RP SEQUENCE OF 157-348 FROM N.A.  
 RC STRAIN=AB320;  
 RX MEDLINE=91326035; PubMed=1650910;  
 RA Itaya M., McKelvin D., Chatterjee S.K., Crouch R.J.;  
 RT "Selective cloning of genes encoding RNase H from *Salmonella*  
 typhimurium, *Saccharomyces cerevisiae* and *Escherichia coli* rnh  
 mutant.";  
 RL Mol. Gen. Genet. 227:438-445(1991).  
 RN [3]  
 RP STRUCTURE BY NMR OF 6-52.  
 RX MEDLINE=99380410; PubMed=10448044;  
 RA Evans S.P., Bycroft M.;  
 RT "NMR structure of the N-terminal domain of *Saccharomyces cerevisiae*  
 RNase H1 reveals a fold with a strong resemblance to the N-terminal  
 domain of ribosomal protein L9.";  
 RL J. Mol. Biol. 291:661-669(1999).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphonooester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit.  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z49939; CA99205.1; -;  
 DR EMBL; X57160; CAA40448.1; -;  
 DR PIR; S57601; S57601.  
 DR PDB; 1OHK; 31-AUG-99.  
 DR GERMOnline; 142909; -;  
 DR SGD; S0004847; RNH1.  
 DR GO; GO:0004523; F:ribonuclease H activity; IDA.  
 DR GO; GO:0007047; P:cell wall organization and biogenesis; IMP.  
 DR InterPro; IPR009027; L9\_N-like.  
 DR Pfam; PF00075; RNaseH.  
 DR Hydrolase; Nuclease; 1.  
 KM Hydrolase; Nuclease; Endonuclease; Magnesium; 3D-structure.  
 FT METAL 193 193 MAGNESIUM (BY SIMILARITY).  
 FT METAL 235 235 MAGNESIUM (BY SIMILARITY).  
 FT METAL 264 264 MAGNESIUM (BY SIMILARITY).  
 FT METAL 338 338 MAGNESIUM (BY SIMILARITY).  
 FT STRAND 6 10  
 FT STRAND 17 21  
 FT STRAND 22 26  
 FT STRAND 38 40  
 FT STRAND 43 50  
 FT HELIX 38 40  
 FT HELIX 43 50  
 FT SEQUENCE 348 AA; 39431 MW; 02403B2329126B90 CRC64;  
 Query Match 16.1%; Score 249.5; DB 1; Length 348;  
 Best Local Similarity 26.0%; Pred. No. 1.1e-14;

Matches 93; Conservative 38; Mismatches 112; Indels 115; Gaps 11;  
 QY 26 GMFVAVRGRKGTGVLPTWNECRQAQVDFPAPARKKATEDEAWAFYRK-----S 74  
 DB 5 GNFYAVRKRGRETIYNTMECKNQVDGYGAIYKKNNSYEQASFLGQPTTISNYGSSTH 64  
 QY 75 ASPPEVBGHEHQGQSEAKPGKRLREPLDGGHESA----- 111  
 DB 65 AGQGVSKPHTTQ-----KVRHRRRPLHYSLSLTSSACSSISANTWTFYVSKSNVNP 117  
 QY 112 -----QPYAKH-----MKRSVPAPP 127  
 DB 118 IESKINNMKDCQAVYHKGRTGTFKKFEDQLAENFTSGMSAHYDLKMNISKESFESKYK 177  
 QY 128 VSRDTFSYMGDFVYVYTDGCGSSNGRRKPRAGIGYVWGPHPLNVG-IRLPGRQTNQRAE 186  
 DB 178 LSSNTW--YKSNVNYCDSSFGNGTSSSRAGGAFEGAPENISEPLSGQTNRRAE 235  
 QY 187 IHAACTAIED--AKTONIKLVLY--TDSMTINGITWVQGWKKNQKMTSAGKEVI-- 239  
 DB 236 IEAVSBAKKIKWEKLTNEKEKVVYQIKTDSYVTKLINDRYMTYDNKKLEGLPNSDLIYP 295  
 QY 240 -----NKEDFVALERLQGMNDIQMNHVPGHSGFGTIGNEADRLAREGAKQ 283  
 DB 296 LVORFVKVKKYIELNKECF---KNNKQFQIEW--VKGHDGDPGENMADFLAKGASR 347  
 RESULT 5  
 RNH\_PASMU  
 ID RNH\_PASMU STANDARD; PRT; 154 AA.  
 AC PE7813;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).  
 GN RNH1 OR RNH OR PM0107.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Pasteurella.  
 NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Em70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;  
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphonooester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE006046; AAK02191.1; -;  
 DR HSSP; P00647; 2RN2.  
 DR HAMAP; MF\_00042; -; 1.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; RNaseH; 1.  
 KM Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 FT METAL 10 10 MAGNESIUM (BY SIMILARITY).  
 FT METAL 48 48 MAGNESIUM (BY SIMILARITY).  
 FT METAL 70 70 MAGNESIUM (BY SIMILARITY).



FT	METAL	134	134	MAGNESIUM (BY SIMILARITY) .
50	SEQUENCE	154 AA;	17556 MW;	37A58196E87661BB CXC64;
	Query Match	15.6%;	Score 240.5;	DB 1; Length 154;
	Beet Local Similarity	34.8%;	Pred. No. 2,5e-14;	
	Matches	54; Conservative	23; Mismatches	49; Indels 29; Gaps 4
Qy	140	VVVYTDGCCSSNGRRRRPRAGIGVYVWPGPHPLNVGIRLPGRO-----	ITNORAEI	187
Db	5	IEIFTDGSCILNP-----	GPQ-----GIGVLLRYKKOHEKOISAGYFLLTNNRMEL	49
Qy	188	HAACAKIAEOAKTONINKLVLTDSMFTNGITNTNWQGGKKQCMKTSAGKEVIVNKEDPAL		247
Db	50	RAVEALNTLKEP--CSYTLHSDSGYMKNGITTKWTFNNKKNNKMSSTCKPVPYNGQDLMQL		107
Qy	248	ERLTQGMHDIQMMWYVPGHSGFIGNEDADRLAREGAK		282
Db	108	DOATQRRHHINQWQKSHGSHLENEICDQDLAKGAE		142

## RESULT 6

ID	RNA TREPA	STANDARD;	PRT;	169 AA.
DT	083372:			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Ribonuclease H (EC 3.1.26.4) (RNAse H).			
GN	RNAse OR TP0353.			
OS	Treponema pallidum.			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.			
OX	NCBI TaxID=160;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Nichols;			
RX	MEDLINE=98332770; PubMed=9665876;			
RA	Fraser C.M., Norris S.U., Weinstein G.M., White O., Sutton G.G.,			
RA	Dodson R., Gwyn M., Hickey E.K., Clayton R., Ketchum K.A.,			
RA	Dodson R., Gwyn M., Hickey E.K., Clayton R., Ketchum K.A.,			
RA	McLoud M.P., Salzberg S., Peterson J.,			
RA	Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,			
RA	McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,			
RA	Hatch B., Horst K., Roberts K., Sandusky M., Wetman J., Smith H.O.,			
RA	Venter J.C.;			
RT	"Complete genome sequence of Treponema pallidum, the syphilis			
RL	spirochete.";			
CC	Science 281:375-388(1998).			
CC	-1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of			
CC	KVA-DNA hybrids specifically (By similarity).			
CC	-1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-			
CC	phosphomonoester.			
CC	-1- COPACITOR: Binds 1 magnesium ion per subunit (By similarity).			
CC	-1- SUBUNIT: Monomer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).			
CC	-1- SIMILARITY: Belongs to the RNase H family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AE001215; AAC65340.1; -			
DR	PIR; G71333; G71333.			
DR	HSSP; P00647; IRMS.			
DR	TIGR; TP0353; -			
DR	HAMAP; MF_00042; -; 1.			
DR	InterPro; IPR002156; RNaseH.			
DR	Pfam; PF00075; rnaesh; 1.			
FM	Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.			
FT	METAL 12 12			
FT	METAL 63 63			
FT	MAGNESIUM (BY SIMILARITY).			
FT	MAGNESIUM (BY SIMILARITY).			

FT	METAL	87	87	MAGNESIUM (BY SIMILARITY).
FT	METAL	151	151	MAGNESIUM (BY SIMILARITY).
SO	SEQUENCE	169 AA;	18184 MM;	164311053632B047 CRC64;
	Query Match	15.5%;	Score 239;	DB 1; Length 169;
	Best Local Similarity	31.5%;	Pred. No. 3.8e-14;	
	Matches	52;	Conservative 25;	Mismatches 58; Indels 30; Gaps 2;
QY	140	VVVYTTDGCSSN-----	-GRRKPRAGIGVYWGPHPLNVI	GRIRLPCR 179
	:			
DB	7	LTTVTDGACLCGPPGCGMAFALVPSDYVPLETGGXAR	EEAAAFRTSG	SGAYP----- 56
QY	180	QTNPRAETIIHAAACKIAEQAKTQNI	INKLVLYTDSMTFINGITNWVQGWKKN	GKTSAGKEVI 239
	:			
DB	57	STNNRMETCAVINIAOEAHGRALAAV	VVVVTTDSQVRKGIQTGMHTWIKNG	WKTAKAPYK 116
QY	240	NKEDFVALLERLTQGMIDIQMMHVP	HSQGIQNEADRIILAREGAKOS	284
	:			
DB	117	NKDIEMEALSLADALSVEMRWVYK	GHAQDPYIELCDRLATIDAPARA	161

## RESULT 7

ID	NAME	STANDARD	PRT	154 AA
AC	08YFR3			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Ribonuclease HI (EC 3.1.26.4) (RNase HI).			
GN	RNA OR EMB11457 OR BR0477.			
OS	Brucella melitensis, and			
OS	Brucella suis.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Brucellaceae; Brucella.			
OX	NCBI_TaxId=29459, 29461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;			
RX	MEDLINE=20020109; Pubmed=11756688;			
RX	Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Muijer C., Los T.,			
RA	Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,			
RA	Jalonenk L., Larsen L., D'Souza M., Bernal A., Mazur M., Goldsman E.,			
RA	Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leeson J.-J.,			
RA	Hasselkorn R., Kyriades N., Overbeek R.;			
RT	"The genome sequence of the facultative intracellular pathogen			
RT	Brucella melitensis.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=B.suis; STRAIN=1330 / Biovar 1;			
RX	MEDLINE=22247741; Pubmed=12271122;			
RA	Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,			
RA	Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beaman M.U.,			
RA	Daugherthy S.C., Debroy R.T., Durkin A.S., Kolonay J.F., Madupu R.,			
RA	Nelson W.C., Ayodeji B., Krati M., Shetty J., Malek J., Van Aken S.B.,			
RA	Redmiller S., Metcalf H., Gill S.R., White O., Salzberg S.L.,			
RA	Hoover D.J., Lindler L.E., Halling S.W., Boyle S.M., Fraser C.N.;			
RT	"The Brucella suis genome reveals fundamental similarities between			
RT	animal and plant pathogens and symbionts.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:13348-13353(2002).			
CC	-1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of			
CC	RNA-DNA hybrids specifically (By similarity).			
CC	-1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-			
CC	phosphomonoester.			
CC	-1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).			
CC	-1- SUBUNIT: Monomer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).			
CC	-1- SIMILARITY: Belongs to the RNase H family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation --			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	way			

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

CC -----  
 CC EMBL; AEO09582; AAL52638.1; -.  
 CC DR EMBL; AEO14357; AAN29420.1; -.  
 CC DR PIR; AC3434; AC3434.  
 CC DR TIGR; BR0477; -.  
 CC DR HAMAP; MF\_00042; -. 1.  
 CC DR InterPro; IPR002156; RNAseH.  
 CC DR Pfam; PF00075; rnaaseh; 1.  
 CC KM Hydroxylase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 CC FT METAL 9 9 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 47 47 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 69 69 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 133 133 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 154 AA; 17172 MW; 3B76C03F52E01012 CRC64;

Query Match 15.4%; Score 238.5; DB 1; Length 154;  
 Best Local Similarity 37.1%; Pred. No. 3.7e-14;  
 Matches 56; Conservative 23; Mismatches 49; Indels 25; Gaps 5;

QY 140 VVVYTDGCGSSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ-----TNQPAETIHA 189  
 Db 4 IEAYTDGACGSGN-----PGPG-----ALIRNNGNKEKLGGEAETTNRMELMA 50

QY 190 ACKAIEQAKTQNIKLVLYTDSMTFTINGITNWQGWKNGKTSAGKEVINKEDEVALER 249  
 Db 51 AIALSLALKEPP--CEVLDLYTDSVYVRDGIQWIEGMRNGMKTPAKKAPVKNALQMLQALD 108

QY 250 LTQGMIDQMMHVPGHSGFTIGNEADRLAREG 280  
 Db 109 ARKAHKVTWMIKSHAGHPENERADELARAAG 139

RESULT 8  
 RNH\_HAEIN STANDARD; PRT; 154 AA.

AC P43807;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HI (EC 3.1.26.4) (RNAse HI).  
 GN RNHA OR RNH OR HI0138.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 NC NCBI\_TaxId=727;  
 RX MEDLINE=7542800;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kariya A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Field C.A., Gocayne J.D.,  
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utechtback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Greth C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT R8.";  
 RL Science 269:496-512(1995).  
 RN [2]  
 RP SEQUENCE OF 84-123 FROM N.A.  
 RA Mizrahi V., Dudding L.R.;  
 CC Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).

CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNAse H family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL, outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

CC -----  
 CC EMBL; U32699; AAC21809.1; -.  
 CC DR EMBL; L11915; AAA25000.1; -.  
 CC DR PIR; C64050; C64050.  
 CC DR HSSP; P00647; 1RBR.  
 CC DR TIGR; HI0138; -.  
 CC DR HAMAP; MF\_00042; -. 1.  
 CC DR InterPro; IPR002156; RNAseH.  
 CC DR Pfam; PF00075; rnaaseh; 1.  
 CC KM Hydroxylase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 CC FT METAL 10 10 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 48 48 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 70 70 MAGNESIUM (BY SIMILARITY).  
 CC FT METAL 134 134 MAGNESIUM (BY SIMILARITY).  
 CC SQ SEQUENCE 154 AA; 17640 MW; 79A413E23DD2BD5A CRC64;

Query Match 15.3%; Score 236.5; DB 1; Length 154;  
 Best Local Similarity 34.8%; Pred. No. 5.6e-14;  
 Matches 54; Conservative 23; Mismatches 49; Indels 29; Gaps 4;

QY 140 VVVYTDGCGSSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ-----TNQPAETIHA 187  
 Db 5 IEFTDGSCLGN-----PGAG-----GIGAVLRKYQHEKTLKSGYFQTTNNRMEI 49

QY 188 HAACKAIEQAKTQNIKLVLYTDSMTFTINGITNWQGWKNGKTSAGKEVINKEDEVALER 247  
 Db 50 RAVIEALNTLKEPCL--ITLYSDQYKNGITKMTINWKKNNKASSGKPVKNQDMLIAL 107

QY 248 ERLTQGMIDQMMHVPGHSGFTIGNEADRLAREGAK 282  
 Db 108 DESIQNHKINIMQVVKHAGHRENEICDELAKKGA 142

RESULT 9  
 RNH\_RHILO STANDARD; PRT; 177 AA.

AC O985M1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribonuclease H (EC 3.1.26.4) (RNAse H).  
 GN RNHA OR MLR7504.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Mesorhizobium.  
 NC NCBI\_TaxId=381;  
 RX MEDLINE=11214968;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RX MEDLINE=MAFF310309;  
 RP SEQUENCE FROM N.A.  
 RA Kaneo T., Nakamura Y., Sato S., Aasamizu E., Kato T., Sasamoto S.,  
 RA Matanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,  
 RA Nouchizaki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.

```
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF003011; BAB53592.1; -
DR HAMAP; MF_000642; -; 1.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; RNaseH; 1.
KW Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 10
FT METAL 48
FT METAL 70
FT METAL 70
FT METAL 134
SQ SEQUENCE 177 AA; 19458 MW; 40BF6ED04C9EB58 CRC64;

Query Match 15.3%; Score 236.5; DB 1; Length 177;
Best Local Similarity 39.6%; Pred. No. 6.7e-14;
Matches 59; Conservative 19; Mismatches 58; Indels 13; Gaps 5;

OY 136 MGDFVVTGGCCSSNGRRKPRAGIGVWPGHPLNVGIR-LPGHQ--TNGRAIHAC 191
DB 1 MSKQVEIFDGCAGSN---PGPG---GKGALIRFGTGTKEISGGEAEFTNNRMELAAI 53

OY 192 KAIEAKTQNTKLVLYTDSMTINGITWVGKKNKTSAGKEVINKPEVALERLT 251
DB 54 SALNALKRP--CTVEHTDTSKYMDGISKWHGKKNKWTADKKPVKNGLMQALDEAN 111

OY 252 QGMDIOMHVPGHSGFTIGNEADRLAREG 280
DB 112 RHKVTWNVKGHAGHTENERADELAREG 140

RESULT 10
RNH_PSEAE STANDARD; PRT; 148 AA.
AC Q91359;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).
GN RNHA OR PA1815.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI.
RX MEDLINE=2043737; PubMed=10984043;
RA Steyer C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RT Nature 406:959-964 (2000).
RT FUNCTION: This enzyme is an endonuclease that degrades the RNA of
RT RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
```

```
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004608; AAG05204.1; -
DR PIR; A83418; A83418.
DR HSSP; P00647; 1RBS.
DR InterPro; IPR002156; RNaseH.
DR HAMAP; MF_000642; -; 1.
DR Pfam; PF00075; RNaseH; 1.
KW Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 12
FT METAL 50
FT METAL 72
FT METAL 72
FT METAL 136
SQ SEQUENCE 148 AA; 16697 MW; 1171AF5A9D267CD8 CRC64;

Query Match 15.1%; Score 233.5; DB 1; Length 148;
Best Local Similarity 35.8%; Pred. No. 9.9e-14;
Matches 54; Conservative 22; Mismatches 56; Indels 19; Gaps 4;

OY 140 VVVTYDGCSSNGRRKPRAGIGVY-----WPGHPLNVGIRLPGRTNGRAIHAACK 192
DB 7 VVIYTDGCKNGPRGCGWGLALLYKGARELWG--GEP-----DTNNRMELMAAIQ 56

OY 193 AIEQAKTQNTKLVLYTDSMTINGITWVGKKNKTSAGKEVINKPEVALERLTQ 252
DB 57 ALAALKRSCPRLR--TTSVYMRGITEWLPWKKRGKWTASKQPVKNADLMQALDEOVA 114

OY 253 QGMDIOMHVPGHSGFTIGNEADRLAREGAKQ 283
DB 115 RHQVEMQVNRGHTGDPGNERRADQLANRGVAE 145

RESULT 11
RNH_NEIMA STANDARD; PRT; 145 AA.
AC Q9UTD9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).
GN RNHA OR NMA1817.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parthill J., Achman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamilton N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RT Nature 404:502-506 (2000).
RT FUNCTION: This enzyme is an endonuclease that degrades the RNA of
RT RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
```

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A162757; CAB85042.1; -.
CC PIR: F81807; F81807.
CC HSP: P00647; 2RN2.
CC HAMAP: MF_00042; -.
CC InterPro: IPR002156; RNaseH.
CC Pfam: PF00075; rnaheH; 1.
CC HydroLase; Nuclease; Endonuclease; Magnesium; Complete proteome.
CC METAL 10 10 MAGNESIUM (BY SIMILARITY).
CC METAL 48 48 MAGNESIUM (BY SIMILARITY).
CC METAL 70 70 MAGNESIUM (BY SIMILARITY).
CC METAL 134 134 MAGNESIUM (BY SIMILARITY).
CC SEQUENCE 145 AA; 16235 MW; 3E2BA6A1317D4171 CRC64;

Query Match 15.0%; Score 232.5; DB 1; Length 145;
Best Local Similarity 34.8%; Pred. No. 1.2e-13;
Matches 55; Conservative 22; Mismatches 56; Indels 25; Gaps 5;

QY 136 MGDFVYVYTDGSSNGRRKPRAGIGYWGPGHPLVNGIRLPGRQ-----TNQRA 185
Db 1 MNQTVLYITDGACKGN---PGAG---GWC-----VLKRGSHKELFGGAQTNNRM 47

QY 186 EHAACKAIEQATQNTNKLVLTDSMFTNGITNNVQGWKKGWTSACKEVINKEDFV 245
Db 48 ELTAIVIEGLSKLRR--CTVITCTDSQYVKNQEMNHWKRWGKMTAKQPVNDDLMK 105

QY 246 ALERLTQGMIDQWMAHPGSGFGINEADRLAREGAQ 283
Db 106 ELDAIVGRHGVSTWVKGHAGHAENERADDLANRGAQ 143

RESULT 12
RNH_NEIMB STANDARD; PRT; 145 AA.
ID RNH_NEIMB
AC QJYES;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).
GN RNHA OR NMB1618.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B.
RC MEDLINE=20175755; PubMed=10710307;
RA Tetselin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Bisen J.A., Kechum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., Deboy O., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,
RA Cotton M.D., Utecherack T.R., Khoult H., Qin H., Vamathavan J.,
RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Frazer C.M., Moxon E.R., Rappoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RT Science 287:1809-1815(2000).
RL -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.

```

```

CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE002512; AAF41970.1; -.
CC PIR: H81061; H81061.
CC HSP: P00647; 2RN2.
CC TIGR: NMB1618; -.
CC HAMAP: MF_00042; -.
CC InterPro: IPR002156; RNaseH.
CC Pfam: PF00075; rnaheH; 1.
CC HydroLase; Nuclease; Endonuclease; Magnesium; Complete proteome.
CC METAL 10 10 MAGNESIUM (BY SIMILARITY).
CC METAL 48 48 MAGNESIUM (BY SIMILARITY).
CC METAL 70 70 MAGNESIUM (BY SIMILARITY).
CC METAL 134 134 MAGNESIUM (BY SIMILARITY).
CC SEQUENCE 145 AA; 16251 MW; 3E2BA6BD1D5B91 CRC64;

Query Match 14.8%; Score 229.5; DB 1; Length 145;
Best Local Similarity 34.2%; Pred. No. 2.2e-13;
Matches 54; Conservative 23; Mismatches 56; Indels 25; Gaps 5;

QY 136 MGDFVYVYTDGSSNGRRKPRAGIGYWGPGHPLVNGIRLPGRQ-----TNQRA 185
Db 1 MNQTVLYITDGACKGN---PGAG---GWC-----VLKRGSHKELFGGAQTNNRM 47

QY 186 EHAACKAIEQATQNTNKLVLTDSMFTNGITNNVQGWKKGWTSACKEVINKEDFV 245
Db 48 ELTAIVIEGLSKLRR--CTVITCTDSQYVKNQEMNHWKRWGKMTAKQPVNDDLMK 105

QY 246 ALERLTQGMIDQWMAHPGSGFGINEADRLAREGAQ 283
Db 106 ELDAIVGRHGVSTWVKGHAGHAENERADDLANRGAQ 143

RESULT 13
RNH_ZYMO STANDARD; PRT; 156 AA.
ID RNH_ZYMO
AC Q69014;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease H (EC 3.1.26.4) (RNase H).
GN RNHA OR RNH.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=542;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 31821 / ZM4 / CP4;
RA Lee J., Jin S., Kang H.S.;
RT "Sequence analysis of a cosmid clone of Zymomonas mobilis ZM4
RT containing alcohol dehydrogenase B gene."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: Belongs to the RNase H family.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF086791; AAC70364.1; --  
DR PIR: T33725; T33725.  
DR HSSP: P00647; 2RN2.  
DR HAMAP: MF\_00042; -; 1.  
DR InterPro: IPR002156; RNaseH.  
DR Pfam: PF00075; RNaseH; 1.  
KW Hydrolase; Nuclease; Endonuclease; Magnesium.  
FT METAL 16 16 MAGNESIUM (BY SIMILARITY).  
FT METAL 54 54 MAGNESIUM (BY SIMILARITY).  
FT METAL 76 76 MAGNESIUM (BY SIMILARITY).  
FT METAL 140 140 MAGNESIUM (BY SIMILARITY).  
SQ SEQUENCE 156 AA; 17251 MW; 5726819A1A2D92A3 CRC64;

Query Match 14.4%; Score 224.5; DB 1; Length 156;  
Best Local Similarity 37.8%; Pred. No. 6.6e-13;  
Matches 56; Conservative 21; Mismatches 54; Indels 17; Gaps 6;

QY 140 VVVTYDGCSSNGRRKPRAGIGVWG-----PGHPLNV-GIRLPGRQTNQRAEIHACKA 193  
DB 11 VVIATDAGACKGN-----PGFG-GWGALIRYQGEKASGSSEN-PTNNRELOQVIERA 61  
QY 194 IEQATQINIKLVLTDSMTFNGITTNVVGKMKKMTSAGKEVINKEDFYALERTLQ 253  
DB 62 LSCLKRP--CQELSTDSKYVMDGLTRWTHGQKKGWMLTAAKKPVKNADLWQLLATLQ 119  
QY 254 MDIQMHVPGHSGFGFNGEADRLAREG 281  
DB 120 HDIAWKVWGAGHAGHPNERADQLASDPA 147

RESULT 14  
RNH\_RHIME STANDARD; PRT; 153 AA.  
AC 092RG0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribonuclease H (EC 3.1.26.4) (RNase H).  
GN RNHA1 OR RNHA1 OR R00914 OR SMC00018.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN SEQUENCE FROM N.A.  
RP STRAIN=1021;  
RA MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Barut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godin T., Goffeau A., Kahn D., Kiss B., Lelure V., Masny D.,  
RA Pohl T., Portetle D., Puehler A., Purnelle B., Ramsparger U.,  
RA Renard C., Thibault P., Vandenberg M., Weidner S., Gallbert F.,  
RA "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
CC RNA-DNA hybrids specifically (By similarity).  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphomonoester.  
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- SIMILARITY: Belongs to the RNase H family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AL591785; CAC45486.1; --  
DR HAMAP: MF\_00042; -; 1.  
DR InterPro: IPR002156; RNaseH.  
DR Pfam: PF00075; RNaseH; 1.  
KW Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).  
FT METAL 47 47 MAGNESIUM (BY SIMILARITY).  
FT METAL 69 69 MAGNESIUM (BY SIMILARITY).  
FT METAL 133 133 MAGNESIUM (BY SIMILARITY).  
SQ SEQUENCE 153 AA; 17178 MW; 3562BE7C672B86D2 CRC64;

Query Match 14.4%; Score 222.5; DB 1; Length 153;  
Best Local Similarity 37.4%; Pred. No. 9.7e-13;  
Matches 55; Conservative 22; Mismatches 53; Indels 17; Gaps 5;

QY 140 VVVTYDGCSSNGRRKPRAGIGVWGPHPLNVG--IRLPGRQ--TNQRAEIHACKA 193  
DB 4 VHIPTDAGACKGN-----PGFG-----GWDALVIRYGVKEKMGGEAETTNREMLLAISA 54  
QY 194 IEQATQINIKLVLTDSMTFNGITTNVVGKMKKMTSAGKEVINKEDFYALERTLQ 253  
DB 55 INALRQP--CEVDLHTDSKYVMDGISKVIHGWKRMKMTGDKPVKNDELWQALDEARR 112  
QY 254 MDIQMHVPGHSGFGFNGEADRLAREG 280  
DB 113 HNVTHVWKGAGHAGHPNERADQLARKG 139

RESULT 15  
RNH\_HAEDU STANDARD; PRT; 153 AA.  
AC 07W15;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).  
GN RNHA1 OR HD1206.  
OS Haemophilus ducreyi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=730;  
RN SEQUENCE FROM N.A.  
RP STRAIN=35000HP / ATCC 700724;  
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.,  
RA "The complete genome sequence of Haemophilus ducreyi."  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
CC RNA-DNA hybrids specifically (By similarity).  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphomonoester.  
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- SIMILARITY: Belongs to the RNase H family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE017154; AAP96053.1; --

```

DR HAMAP: ME-00042; -, 1.
DR InterPro: IPR002156; RNaseH.
DR Pfam: PF00075; rnaaseh; 1.
KM Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).
FT METAL 47 47 MAGNESIUM (BY SIMILARITY).
FT METAL 69 69 MAGNESIUM (BY SIMILARITY).
FT METAL 133 133 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 153 AA; 17453 MW; 390C8F87B24F1DD1 CRC64;

Query Match 14.3%; Score 221.5; DB 1; Length 153;
Best Local Similarity 35.1%; Pred. No. 1.2e-12;
Matches 54; Conservative 19; Mismatches 52; Indels 29; Gaps 4;

QY 140 VVVYTDGCGSSNGRRPRAGIGVYWGPHLVNGIRLPGRQ-----TNCPRAET 187
DB 4 VNIFTDGSCLGNP-----GFG-----GIGVAVLRNQHOKKVSQGVQFTNNMEL 48

QY 188 HAACKAIEQAKTQNTKLVLYTDSMTFINGITNNVQKKGKMTSAGKEVINKEDFVAL 247
DB 49 RAYIEGLSMRK--EACVVTLYSDSQYMKNGITTKVIFPKKXSNWKTANGKAVKXKDWILL 106

QY 248 ERLTQGMIDQMHPGHSFGFNGEADRLAREGA 281
DB 107 DEKLIHYIEKMKVKGSHGHEHCDELAKLGA 140

RESULT 16
RNH_AGRTS STANDARD; PRT; 146 AA.
AC QGUAH7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Ribonuclease H (EC 3.1.26.4) (RNAse H).
GN RNHA OR ATTU076 OR AGR C.1417.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiales; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Secudal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Decherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung W., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323 (2001).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quirillo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Bpp A., Liu F.,
RA Woliam K., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Plangan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328 (2001).
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).

```

```

CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE009044; AL4192.1; -.
DR EMBL: AE008010; AAK8585.1; -.
DR PIR: AB2672; AB2672.
DR PIR: H97453; H97453.
DR HAMAP: ME-00042; -, 1.
DR InterPro: IPR002156; RNaseH.
DR Pfam: PF00075; rnaaseh; 1.
KM Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).
FT METAL 47 47 MAGNESIUM (BY SIMILARITY).
FT METAL 69 69 MAGNESIUM (BY SIMILARITY).
FT METAL 133 133 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 146 AA; 16261 MW; CA90796A79057155 CRC64;

Query Match 14.3%; Score 220.5; DB 1; Length 146;
Best Local Similarity 40.1%; Pred. No. 1.4e-12;
Matches 59; Conservative 17; Mismatches 54; Indels 17; Gaps 5;

QY 140 VVVYTDGCGSSNGRRPRAGIGVYWGPHLVNG---IRLPGRQ---TNCPRAIHAACKA 193
DB 4 VDIFTGACSGN-----GFGP-----GMAVLRVGETEKEISGGEADPTNNRMELLAISA 54

QY 194 IEQAKTQNTKLVLYTDSMTFINGITNNVQKKGKMTSAGKEVINKEDFVALERTQG 253
DB 55 LNLAKSP--CEVLLYDSAYVADGITTKVIFGKKGKMTADKPVQNVLELMALAAQER 112

QY 254 MDIQMHWHPGHSFGFNGEADRLAREGA 280
DB 113 HKVTLHWVKGHAGHPENRADERALARKG 139

RESULT 17
RNH_CAUCR STANDARD; PRT; 149 AA.
AC Q9A341;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Ribonuclease HI (EC 3.1.26.4) (RNAse HI).
GN RNHA OR CC3365.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Kouri H., Shetty J., Berry K.,
RA Uettersack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.

```



```
CC -1- COPACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity). (Potential).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB005998; AAK25327.1; -.
DR PIR; C87666; C87666.
DR HSP; P00647; IRBR.
DR TIGR; CC3365; -.
DR HAMAP; MF_00042; -.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; rnaseh; 1.
DR Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 10 MAGNESIUM (BY SIMILARITY).
FT METAL 48 48 MAGNESIUM (BY SIMILARITY).
FT METAL 70 70 MAGNESIUM (BY SIMILARITY).
FT METAL 134 134 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 149 AA; 16674 MW; CB2271A33782435C CRC64;

Query Match 14.2%; Score 219.5; DB 1; Length 149;
Best Local Similarity 36.1%; Pred. No. 1.7e-12;
Matches 56; Conservative 21; Mismatches 53; Indels 25; Gaps 6;

QY 140 VVVVYDGGCCSSNGRRKPRAGIGVWGPGRPLVNGVRL-----PGRQNOAEIHAACK 192
DB 5 VVITYDGAACKGN-----GPGG-----GWGALIFYGDKKEICGGEPC-TNNNMEIWLAAIQ 54
QY 193 AIEQAKTQINK---LVLTDSMFTINGITNNVQCKKMGKTSAGKEYINKEDFVALER 249
DB 55 ALLEL-----LNPFCVELHTDSQYVWKGIGQEWIRGKAKGWKTADSPKVDLMWRDLA 109
QY 250 LTQGMIDQWVHVGSHGFIQNEADRLAREGAKOS 284
DB 110 AARAHVDVWRVWVKGHAGHPLNERADLANEGRLQA 144

RESULT 18
RNH_PSESM STANDARD; PRT; 150 AA.
AC 0877T0;
RX MEDLINE=22834015; PubMed=12928499;
RA Buell C.R., Joarder V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beaman M.J., Haft D.H.,
RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Uterback T.,
RA Van Aken S.E., Feldblum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
RA Lazarowicz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collier A.;
RA "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RA Proc. Natl. Acad. Sci. U.S.A. 100:10161-10166(2003).
```

```
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- COPACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB016869; AAO57181.1; -.
DR TIGR; PSP03712; -.
DR HAMAP; MF_00042; -.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; rnaseh; 1.
DR Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 10 MAGNESIUM (BY SIMILARITY).
FT METAL 48 48 MAGNESIUM (BY SIMILARITY).
FT METAL 70 70 MAGNESIUM (BY SIMILARITY).
FT METAL 134 134 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 150 AA; 16782 MW; 4ED1FC436407C77C CRC64;

Query Match 14.2%; Score 219.5; DB 1; Length 150;
Best Local Similarity 33.1%; Pred. No. 1.7e-12;
Matches 53; Conservative 20; Mismatches 58; Indels 29; Gaps 5;

QY 136 MGDFVVVYDGGCCSSNGRRKPRAGIGVWGP-----GRLVNGVRLPGRQNO 183
DB 1 MSDSYELFTDGAACKGN-----PQPG-----GWGALIVCKGVEKEIWCGEANT-----TNN 45
QY 184 RAETHAACKALQEQATQONINKLVLTDSMFTINGITNNVQCKKMGKTSAGKEYINKED 243
DB 46 RMELTGALRGLEELRP--CEVTLVTDQYVWKGIGTEWVNNKKGWKTADSPKVNADL 103
QY 244 FVALERLTQGMIDQWVHVGSHGFIQNEADRLAREGAKQ 283
DB 104 WQLDQVSRHTVVKWQWRGHTGHGNERADQLANRGVDE 143

RESULT 19
RNH_PSEPK STANDARD; PRT; 148 AA.
AC 088F5;
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beaman M., Deboy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple B., Scanlan D., Tran K., Moazzez A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Medler H.,
RA Lauber U., Stepanic D., Hohnselt J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RA "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
```

```

RL Environ. Microbiol. 4:799-806(2002).
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC -1- RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC -1- phosphononucleoside.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE016789; AAM69724.1; -.
DR TIGR, PA142; -.
DR HAMAP, MF 00042; -1.
DR InterPro, IPR002156; RNaseH.
DR Pfam, PF00075; rnaaseH; 1.
KM Hydroxylase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 10 10 MAGNESIUM (BY SIMILARITY).
FT METAL 48 48 MAGNESIUM (BY SIMILARITY).
FT METAL 70 70 MAGNESIUM (BY SIMILARITY).
FT METAL 134 134 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 148 AA; 16913 MW; CB869BD513B8953 CRC64;

Query Match 14.1%; Score 218.5; DB 1; Length 148;
Best Local Similarity 35.5%; Pred. No. 2,1e-12;
Matches 54; Conservative 21; Mismatches 64; Indels 13; Gaps 4;

QY 136 MGDFVVVYTDGSSSSNGRRPRAG-----IGYMGPRGPLVNGIRLPQRTNORAEIYAC 191
DB 1 MSDSVEFTGACKGN-----PGRGMGVLMITKVEVELNGER---ETTNNKELMAAI 53

QY 192 KAEQAKTONINKLVLTYSMTFTINGITNNVQKXKXGKTSAGKEVINKEDFVALERT 251
DB 54 QGLMSLRKRE--CEVVLTTDSQYWKGINEMVNMKKGKMTAKAPEKXNDLWQDDEGV 111

QY 252 QGMDIQMHVPGHSGFTGNEBDRILAREGAQ 283
DB 112 NRHKVTKWVRGHIHGPNERADQLANRGVDE 143

RESULT 20
RNH_XANCP STANDARD; PRT; 150 AA.
AC Q8PBX8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease H (EC 3.1.26.4) (RNase H).
GN RNHA OR XCC0987.
OS Xanthomonas campestris (gv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxId=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Camavan F., Cardozo J., Chambergo F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorty H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,

```

```

RA Martins E.C., Meidania J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitzajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC -1- RNA-DNA hybrids specifically (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC -1- phosphononucleoside.
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the RNase H family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE012196; AAM40289.1; -.
DR HAMAP, MF 00042; -1.
DR InterPro, IPR002156; RNaseH.
DR Pfam, PF00075; rnaaseH; 1.
KM Hydroxylase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT METAL 9 9 MAGNESIUM (BY SIMILARITY).
FT METAL 47 47 MAGNESIUM (BY SIMILARITY).
FT METAL 69 69 MAGNESIUM (BY SIMILARITY).
FT METAL 133 133 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 150 AA; 16743 MW; DE84CDB07D0B172 CRC64;

Query Match 14.1%; Score 217.5; DB 1; Length 150;
Best Local Similarity 35.5%; Pred. No. 2,6e-12;
Matches 55; Conservative 15; Mismatches 54; Indels 31; Gaps 4;

QY 140 VVVYTDGSSS-----NGRRPRAGIGYMGPRGPLVNGIRLPQRTNORAE 186
DB 4 IEVHTDSSCLGNPGRGMALLRYNGRKEKLG-----GE-----ANSTNNRHE 47

QY 187 IHAACKAIEQAKTONINKLVLTYSMTFTINGITNNVQKXKXGKTSAGKEVINKEDFYA 246
DB 48 IMAAIMALELTTP--CGILLHTDSQYVQGITEMWPGVNRGKXGSDPVYNNRLMER 105

QY 247 LERLTQGMDIQMHVPGHSGFTGNEBDRILAREGA 281
DB 106 LHAATQRHSIEWRWKGNHNDPNERHVDYLARQA 140

RESULT 21
RNH_RALSO STANDARD; PRT; 151 AA.
AC Q8XZ91;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease H (EC 3.1.26.4) (RNase H).
GN RNHA OR RSC1513 OR RS03791.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxId=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Gentin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Atlat M., Billault A., Brottier P., Camus J.C., Catolico L.,

```



RA Chandler M., Choiane N., Claudel-Renard C., Cunac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Siegler P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissbach J., Boucher C.A.,  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"  
 RL Nature 415:497-502(2002).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -1- COFACTOR: Magnesium (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AL646065; CAD15215.1; -.  
 DR HAMAP: MF\_00042; -; 1.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 KM Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 FT METAL 9  
 FT METAL 9  
 FT METAL 47 47  
 FT METAL 69 69  
 FT METAL 133 133  
 SQ SEQUENCE 151 AA; 16598 MW; A64CDD79C7F98143 CRC64;  
 Query Match 14.0%; Score 216.5; DB 1; Length 151;  
 Best Local Similarity 35.8%; Pred. No. 3.3e-12;  
 Matches 54; Conservative 20; Mismatches 56; Indels 21; Gaps 6;  
 QY 140 VVVVYDGGCCSSNGRRKPRAGIGVYNGP-----GHPLVNIGIRLPGRQ---TNGRAEIHAAK 191  
 DB 4 VTVVSDGACKGN-----PGLG-GWGTVLVSGGHEK---ELPGGAVTNNRMEIMAVT 52  
 QY 192 KAIEOAKTONINKLVYTDSTMTFINGITNWQGWKKGKTSAGKEVINKEDEVALERLTQ 251  
 DB 53 EAFRLAKRP--CRVAVYTDSTQVQVQKISGEMLAGWARGKTKADKPKVKNDDMLRTIDELV 110  
 QY 252 QGMDIQMHVPGHSGPTIGNEADRLAREGAK 282  
 DB 111 VTHEVSWHWKGGHAGHPGERADALANKGVE 141  
 RESULT 22  
 RNH\_MYCSM STANDARD; PRT; 159 AA.  
 AC 007705;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribonuclease H (EC 3.1.26.4) (RNase H).  
 GN RNHA.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OK NCBI\_TaxID=1772;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=LR22;  
 RC Daves S.S., Crouch R.J., Morris S.L., Mizrahi V.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBD databases.  
 [2]  
 SEQUENCE OF 44-110 FROM N.A.  
 RC STRAIN=LR22;  
 RX MEDLINE=94124016; PubMed=8294019;

RA Mizrahi V., Huberts P., Daves S.S., Dudding L.R.;  
 RT "A PCR method for the sequence analysis of the *gyrA*, *polA* and *rnha*  
 RT gene segments from mycobacteria,"  
 RL Gene 136:287-290(1993).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC RNA-DNA hybrids specifically.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium-ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U20115; AAA62124.1; -.  
 DR HSSP: P00647; IGOA.  
 DR HAMAP: MF\_00042; -; 1.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 KM Hydrolase; Nuclease; Endonuclease; Magnesium.  
 FT METAL 11 11  
 FT METAL 50 50  
 FT METAL 72 72  
 FT METAL 136 136  
 FT METAL 45 45  
 FT CONFLICT 49 49 M -> L (IN REF. 2).  
 SQ SEQUENCE 159 AA; 17529 MW; EF0875F3B6D92BF CRC64;  
 Query Match 14.0%; Score 216; DB 1; Length 159;  
 Best Local Similarity 36.8%; Pred. No. 3.8e-12;  
 Matches 56; Conservative 24; Mismatches 58; Indels 14; Gaps 6;  
 QY 138 DRVVYTTGGCCSSNGRRKPRAGIGVYNGP--GHPLVNIGIRLPGR---QTNGAEIHAAK 192  
 DB 4 DVIHHTDGGC-----RPNPGG--GWGAVLHHRHVEEMFGGAAYTSNNMELTAPIM 56  
 QY 193 AIEOAKTONINKLVYTDSTMTFINGITNWQGWKKGKTSAGKEVINKEDEVALERLTQ 252  
 DB 57 ALE-ALTBPVT-VHLYTSTYRANGITKVLGHEKNGMTAKQPVKVDLMQRLQACA 114  
 QY 253 GMDIQMHVPGHSGPTIGNEADRLAREGAKOS 284  
 DB 115 RHQVEMFWKGGHSGIGDNEIDLATRLQLEA 146  
 RESULT 23  
 RNH\_YERPE STANDARD; PRT; 154 AA.  
 AC 08230;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).  
 GN RNHA OR YPO1081 OR Y3095.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersiniaceae; Yersinia.  
 OK NCBI\_TaxID=632;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=CO-92 / Biovar Orientalis;  
 RC MEDLINE=21470413; PubMed=11586360;  
 RX Parkhill J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Chuchter C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Kariyshev A.V.,  
 RA Leather S., Mould S., Oyston P.C.F., Quail M.A., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.,  
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague,"  
 RT Nature 413:523-527(2001).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM5 / Biovar Mediaevalis;  
 RX MEDLINE=22137863; PubMed=12142430;  
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
 RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,  
 RA Perry R.D.,  
 RT "Genome sequence of *Yersinia pestis* KIM,"  
 RT J. Bacteriol. 184:4601-4611(2002).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC -1- RNA-DNA hydride specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC -1- phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: AJ141416; CAC89924.1; -;  
 DR EMBL: AE013910; AAM8645.1; -;  
 DR PIR: A10132; A10132.  
 DR HAMAP: MF\_00042; -; 1.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 KW Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 FT METAL 10 10 MAGNESIUM (BY SIMILARITY).  
 FT METAL 48 48 MAGNESIUM (BY SIMILARITY).  
 FT METAL 70 70 MAGNESIUM (BY SIMILARITY).  
 FT METAL 134 134 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 154 AA; 17463 MW; 3AB40DE24F3CF507 CRC64;  
 Query Match 13.8%; Score 214.5; DB 1; Length 154;  
 Best Local Similarity 35.7%; Pred. No. 5e-12;  
 Matches 55; Conservative 16; Mismatches 54; Indels 29; Gaps 4;  
 QY 140 VVVYTDGSSNGRRKPRAGIGVWGPGLNVLGIRLPGRO-----TNGRAEI 187  
 DB 5 VEFTDGSCLGNP-----GPG--GYGALRYKHGHEKTFAGYLLTNNRMEL 49  
 QY 188 HAACKAIEQAKTONIKLVLTDSMTFTINGITNWQVKKNGKTSAGKVINKEPVAL 247  
 DB 50 MAIVALEALTPS--CEVLTSTDSQYVRGIGITQWIMHWKKRGWKTDKREVRVNDLMQRL 107  
 QY 248 ERLTQGMIDQGMVHPGHSFGFIGNSEADRLAREGA 281  
 DB 108 DLAIQSHITQEMVKGHAGHPENRCDELARQGA 141  
 RESULT 24  
 RNH\_VIBCH STANDARD; PRT; 156 AA.  
 AC 09KEX8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).  
 GN RNHA OR VC2234.  
 OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OC NCBI\_TaxId=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Sisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Emdon-Lea M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Uetembach T., Fleischmann R.D., Nieman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.,  
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
 RT cholerae,"  
 RT Nature 406:477-483(2000).  
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
 CC -1- RNA-DNA hydride specifically (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC -1- phosphomonoester.  
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the RNase H family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: AE004295; AAF95378.1; -;  
 DR PIR: G82101; G82101.  
 DR HESP: P00647; 2RN2.  
 DR TIGR: VC2234; -; 1.  
 DR HAMAP: MF\_00042; -; 1.  
 DR InterPro: IPR002156; RNaseH.  
 DR Pfam: PF00075; rnaseh; 1.  
 KW Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
 FT METAL 10 10 MAGNESIUM (BY SIMILARITY).  
 FT METAL 48 48 MAGNESIUM (BY SIMILARITY).  
 FT METAL 70 70 MAGNESIUM (BY SIMILARITY).  
 FT METAL 134 134 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 156 AA; 17953 MW; 6197AD5E65460D96 CRC64;  
 Query Match 13.8%; Score 214; DB 1; Length 156;  
 Best Local Similarity 34.1%; Pred. No. 5.6e-12;  
 Matches 56; Conservative 21; Mismatches 57; Indels 30; Gaps 5;  
 QY 136 MEDFVVVYTDGSSNGRRKPRAGIGVWGP-----HPLNVLGIRLPGRQTNQ 183  
 DB 1 MKQVEIFPDGSLGNP-----GPGGYGVWAKYQVEKTLAIGYRL--TTNN 45  
 QY 184 RAEIHAAKRAIEQAKTONIKLVLTDSMTFTINGITNWQVKKNGKTSAGKVINKEPVAL 243  
 DB 46 RMEMLAAMVALQALKEP--CRVILTTDSQYVRGIGITQWIMHWKKRGWKTDKREVRVNDLMQRL 103  
 QY 244 FVALERLTQGMIDQGMVHPGHSFGFIGNSEADRLAREGAQO--SED 286  
 DB 104 WQALDKETARHQVEMRWVKGHAGHRENMCDDELARQGAENPTED 147  
 RESULT 25  
 RNH\_VIBVU STANDARD; PRT; 155 AA.  
 AC 08DBD5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).  
 DE Ribonuclease HI

GN RNHA OR VV11886.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
RA Choy H.E.,  
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of  
CC RNA-DNA hybrids specifically (By similarity).  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphomonoester.  
CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
CC -1- SUBUNIT: Monomer (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- SIMILARITY: Belongs to the RNase H family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
CC EMBL; AE016803; AA010288.1; -.  
DR HAMAP; MF\_00042; -; 1.  
DR InterPro; IPR002156; RNaseH.  
DR Pfam; PF000075; rnaseh; 1.  
KM Hydroxylase; Nuclease; Endonuclease; Magnesium; Complete proteome.  
FT METAL 10 10 MAGNESIUM (By similarity).  
FT METAL 48 48 MAGNESIUM (By similarity).  
FT METAL 70 70 MAGNESIUM (By similarity).  
FT METAL 134 134 MAGNESIUM (By similarity).  
SQ SEQUENCE 155 AA; 17647 MW; E0F8F026F87EC079 CRC64;  
-----  
Query Match 13.6%; Score 209.5; DB 1; Length 155;  
Best Local Similarity 39.0%; Pred. No. 1.4e-11;  
Matches 57; Conservative 16; Mismatches 62; Indels 11; Gaps 5;  
-----  
QY 140 VVVYTDGCCSSNGRRKPRAGIV--YMGPHPLNVGIRLPGRQTNQRAEITHACKATIQ 196  
DB 5 VEIPTDGSCLGN--PGPGYGVVLRYQVETKLAQYRL--TTNNMEMMAATTVAL-Q 57  
QY 197 AKTONINKLVLYTDSMTFTINGITNNVQGWKKNQKMTSAGKEVINKEDEPVALERLTQMDI 256  
DB 58 ALKEPCN-VILTTDSQYVRQGITQYIHNNKKGKWTADKKPYKNDLWQALDKETTRHTI 116  
QY 257 QMMHVPGHSGFIGNEDRLAREGAK 282  
DB 117 DWRWYVGHAGHRENMEDLARAAAE 142

Search completed: October 7, 2004, 08:05:41  
Job time : 25 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:02:56 : Search time 41 Seconds

(without alignments)  
670.995 Million cell updates/sec

Title: US-10-054313-1

Perfect score: 1546

Sequence: 1 MSMLFLAHRVALALPCR.....FIGNEADRLAREGAKQSD 286

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : 1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	440.5	28.5	369	2	T16506
2	422	27.3	333	2	UC5787
3	323.5	20.9	264	2	T40244
4	319.5	20.7	264	2	T43641
5	299.5	19.4	494	2	A46831
6	249.5	16.1	348	2	S57601
7	239	15.5	169	2	G71333
8	238.5	15.4	154	2	AC3434
9	228	15.4	302	2	AF1824
10	236.5	15.3	154	2	C64050
11	233.5	15.1	148	2	AB3418
12	232.5	15.0	145	2	P81807
13	229.5	14.8	145	2	H81061
14	224.5	14.5	156	2	T33725
15	220.5	14.3	146	2	H97453
16	220.5	14.3	149	2	C87666
17	219.5	14.2	149	2	AB1032
18	214.5	13.9	154	2	G82101
19	214	13.8	156	2	NRECH
20	208.5	13.5	155	1	NRECH
21	208.5	13.5	155	2	B90655
22	208.5	13.5	155	2	B85506
23	206.5	13.4	155	2	S21659
24	206.5	13.4	155	2	AB0534
25	194.5	12.6	153	2	F82591
26	193	12.5	140	2	C97214
27	191.5	12.4	240	2	T18916
28	189	12.2	235	2	T35324
29	188.5	12.2	152	2	E71632

30	183.5	11.9	155	2	D97838	ribonuclease H (EC
31	180.5	11.7	697	2	A26132	gag-abl-pol polypr
32	174.5	11.3	160	2	S76609	hypothetical prote
33	173	11.2	315	2	T34502	hypothetical prote
34	171.5	11.1	161	2	A40593	ribonuclease H - B
35	171	11.1	166	2	A42673	ribonuclease H (EC
36	169.5	11.0	146	2	D82315	probable ribonucle
37	163	10.5	1189	1	GNNVM7	HIV-1 retropepsin
38	160	10.3	1165	1	GNNJGL	HIV-1 retropepsin
39	158	10.2	223	2	B72269	ribonuclease H-rel
40	157	10.2	179	2	G75462	ribonuclease H - D
41	155	10.0	476	2	S04842	pol polyproteins -
42	155	10.0	1204	2	S35475	pol polyproteins -
43	155	10.0	1784	2	T10532	gag-pol polyprotei
44	152	9.8	1204	2	S70393	pol polyproteins -
45	150	9.7	1199	1	GNNVM1	HIV-1 retropepsin
46	148.5	9.6	1035	1	GNNJGG	HIV-1 retropepsin
47	147	9.5	1046	1	GNNVCE	pol polyproteins -
48	145	9.4	1196	1	GNNMVR	HIV-1 retropepsin
49	143	9.2	843	1	GNNVMK	pol polyproteins -
50	143	9.2	1196	1	GNNMVG	HIV-1 retropepsin
51	142.5	9.2	1124	2	S23820	pol polyproteins -
52	142	9.2	146	2	B81260	ribonuclease H (EC
53	140.5	9.1	1124	1	GNNJFP	HIV-1 retropepsin
54	138.5	9.0	1032	2	S12153	pol polyproteins -
55	138	8.9	133	2	T52527	probable ribonucle
56	138	8.9	581	2	A42743	pol polyproteins -
57	135.5	8.8	1124	2	B4557	HIV-1 retropepsin
58	134.5	8.7	1055	1	GNNJST	HIV-1 retropepsin
59	134	8.7	1101	1	B45390	HIV-1 retropepsin
60	133	8.6	1086	1	B46335	HIV-1 retropepsin
61	133	8.6	1101	1	GNNJVS	HIV-1 retropepsin
62	132.5	8.6	1034	1	GNNJCA	HIV-1 retropepsin
63	130	8.4	157	2	AC2947	ribonuclease H (lm
64	130	8.4	157	2	G96335	ribonuclease H PAL
65	129.5	8.4	1053	1	GNNJBT	HIV-1 retropepsin
66	127	8.2	1087	2	JQ1162	pol protein - Maed
67	125.5	8.1	648	2	T23621	hypothetical prote
68	125.5	8.1	1036	1	GNNJG2	HIV-1 retropepsin
69	125.5	8.1	1056	1	GNNJG3	HIV-1 retropepsin
70	125	8.1	1189	2	T30319	liam-aa1 retrotran
71	124.5	8.1	1055	2	S53092	retrovirus-related
72	124	8.0	848	4	A44282	ribonuclease H-rel
73	121	7.8	196	2	G83757	retrovirus-related
74	119.5	7.7	867	1	GNNMIA	pol protein - siml
75	117.5	7.6	958	2	S15566	pol protein - siml
76	117.5	7.6	1161	2	S18738	HIV-1 retropepsin
77	117	7.6	1109	1	B45345	HIV-1 retropepsin
78	116.5	7.5	1061	1	GNNJG4	HIV-1 retropepsin
79	115.5	7.5	886	2	T28114	hypothetical prote
80	115.5	7.5	404	2	GNNJSP	pol polyproteins -
81	114.5	7.4	1027	1	GNNJST	HIV-1 retropepsin
82	112.5	7.3	522	2	JA0072	hypothetical prote
83	112.5	7.3	1058	2	S08436	pol polyproteins -
84	111.5	7.2	143	2	E64602	ribonuclease H - H
85	109	7.1	741	2	T24755	hypothetical prote
86	108.5	7.0	1019	2	T11560	pol polyproteins -
87	108.5	7.0	1039	2	S46347	pol polyproteins -
88	108	7.0	1157	1	GNNJLK	pol polyproteins -
89	107	6.9	902	2	T01668	pol polyproteins -
90	105.5	6.8	1145	1	GNNJEV	HIV-1 retropepsin
91	104.5	6.8	243	2	D71911	ribonuclease H-rel
92	104.5	6.8	245	2	A97129	ribonuclease H (f
93	104.5	6.8	1146	1	GNNJEW	HIV-1 retropepsin
94	104.5	6.8	1146	1	GNNJ22	HIV-1 retropepsin
95	104	6.7	1012	1	GNNMVL	HIV-1 retropepsin
96	103.5	6.7	1295	2	T30528	reverse transcript
97	103	6.7	426	2	S37765	hypothetical prote
98	102	6.6	1003	1	GNNMVL	HIV-1 retropepsin
99	101.5	6.6	656	2	S30484	pol polyproteins -
100	101	6.5	814	1	GNNMIP	retrovirus-related



T43641  
ribonuclease H1 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
C:Accession: T43641  
R:Torawa, Y.; Crouch, R.J.  
submitted to the EMBL Data Library, February 1998  
A:Description: Genomic and cDNA sequences of Schizosaccharomyces pombe Ribonuclease H1.  
A:Reference number: 222591  
A:Accession: T43641  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <TOZ>  
A:Cross-references: EMBL:AF048992; PIDN:AA034366.1  
A:Experimental source: strain 912  
A:Gene: rnh1  
A:Introns: 23/2

Query Match 20.7%; Score 319.5; DB 2; Length 264;  
Best Local Similarity 29.3%; Pred. No. 9,6e-21;  
Matches 83; Conservative 42; Mismatches 105; Indels 53; Gaps 5;  
DB 28 FYAVRGRKTYGVFLTWNECRAQVDRFPAPAFKKFATEDEMAFVR----- 72  
8 YYAVARGNTGISTWDEASDQVKGCGNRKYPSEYEAQFCEKTESRYSSSGPPYR 67  
QY 73 -----KSAPSEVSEGHENQOE-SEAKPGKRLREPLDGDGESAQYAKMKPS 121  
DB 68 STTSYGSPSSSSSNYSARHSDKYRKIKISRSYTEKDIEISNDTHKS----- 117  
QY 122 VEPAPVSDRTFSYMGDPFVYVTTDCCSSNGRRKPRAGICGVYWGPGHPLNVGIRLPG-RQ 180  
DB 118 -----IACSDROVVAEGSSSLRNGKKGAVAGGVFFGNDPNNISVPLAGEEQ 165  
QY 181 TNGRAEIHAAACKAIEQAKTQNTKLVLYTDSMTTNGITNTWYQWKGKMGKTSAGKEVIN 240  
DB 166 TNNRAELQAILALENTS-----GDLTRSDSNYSIKSLTTLWPKMKKNDPKTNSQPVKN 221  
QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQ 283  
DB 222 LDLINRASLDMSDRNVSLSEYVKGSHSTDYGNQADMLARRGASE 264

RESULT 5  
A48683  
ribonuclease H1 homolog (EC 3.1.-.-) - Crithidia fasciculata  
C:Species: Crithidia fasciculata  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 21-Jul-2000  
C:Accession: A48683  
R:Campbell, A.G.; Ray, D.S.  
Proc. Natl. Acad. Sci. U.S.A. 90, 9350-9354, 1993  
A:Title: Functional complementation of an Escherichia coli ribonuclease H mutation by a  
A:Reference number: A48683; MUID:94022373; PMID:8415705  
A:Accession: A48683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-494 <CAM>  
A:Cross-references: GB:L18916; MID:G5776548; PIDN:AAA03546.1; PID:G310977  
C:Keywords: hydrolase

Query Match 19.4%; Score 299.5; DB 2; Length 494;  
Best Local Similarity 27.4%; Pred. No. 1.2e-18;  
Matches 98; Conservative 44; Mismatches 121; Indels 95; Gaps 10;  
DB 10 RVALAALPCRRSGRGFMFYAVRGRKTYGVFLTWNECRAQVDRFPAPAFKKFATEDEMAFV 69  
DB 141 RSCAPPE---ASRMKPSFYVAVVGRGIGYSTWDCCSQVGFSAVYKSFRTTSEARA 197  
QY 70 FVRKASPEVSEGHENQOE-----SEAKPGKRLREPLDGDG---HESQOPYAKM 118  
DB 198 YL--TAHPARSGLEKSDRGDAASLSALSEPOVGLRRSRAAEASLYVVEADPAQPTLROR 255

QY 119 KPSVEP--APPVSDTFSYMGDPFVYVTTDCCSSNGRRKPRAGIGVWG---PGHPLNVG 173  
DB 256 VEEVPSGAADVQRESEVP---QVYVYDGCASHNGTTRKARAGYGGFPGSTSDSNFSLP 312  
QY 174 IRLPGRTNORAEIHAAKAEQA-----KTQNTKLVLY 208  
DB 313 VITBAQNNNGEMKAVIHICIVQGVADGVPALGTSHCVPEPWELSELPOPLRLRLVLY 372  
QY 209 TDSMTTNGITNTWYQWKGKMGKTSAGKEVINKEPVALERL----- 250  
DB 373 TDSRYVIDGLTRYALKWANGFKLASKPEPVQDMLRQLIRLDAVNTRYAEQOHHMAAT 432  
QY 251 -----TQMDIQMHVPGHSGFTIGNEADRLAREGAK 282  
DB 433 CSHASTRVPAASQSKRFHTNTRNDETGEILR--HVGHSHNDYGNEMADVLAVAGAR 488

RESULT 6  
S57601  
ribonuclease H (EC 3.1.26.4) - Yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YN959.16; protein YMR234W  
C:Species: Saccharomyces cerevisiae  
C:Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 06-Feb-1998  
C:Accession: S57601; S16816  
R:Skellton, J.; Churcher, C.M.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S57587  
A:Accession: S57601  
A:Molecule type: DNA  
A:Residues: 1-348 <SKX>  
A:Cross-references: EMBL:249939; MID:g887599; PID:g887615; MIPS:YMR234W  
A:Experimental source: strain AB972  
R:Itaya, M.; McKelvin, D.; Chatterjee, S.K.; Crouch, R.J.  
Mol. Gen. Genet. 227, 438-445, 1991  
A:Title: Selective cloning of genes encoding RNase H from Salmonella typhimurium, Sacchar  
A:Reference number: S16816; MUID:91326035; PMID:1650910  
A:Accession: S16816  
A:Molecule type: DNA  
A:Residues: 183-348 <ITA>  
A:Cross-references: EMBL:X57160  
C:Genetics:  
A:Gene: SGD:RNH1  
A:Cross-references: SGD:S004847; MIPS:YMR234W  
A:Map position: 13R  
C:Keywords: hydrolase

Query Match 16.1%; Score 249.5; DB 2; Length 348;  
Best Local Similarity 26.0%; Pred. No. 2.2e-14;  
Matches 93; Conservative 38; Mismatches 112; Indels 115; Gaps 11;  
DB 26 GMFYVVRGRKTYGVFLTWNECRAQVDRFPAPAFKKFATEDEMAFVR-----S 74  
DB 5 GMFYVVRGRKREYINTWNECQVQDVGAGLYKKFENYEQAKSFLGQPNNTSNYSSSTH 64  
QY 75 ASPEVSEGHENQOESEAKPGKRLREPLDGDGESA----- 111  
DB 65 AGGVSKPHHTT-----KVRHRRRPLHSLTSSACSSLSAANTTFFSYKSNVNV 117  
QY 112 -----QPYAKH-----MKPSVEPAP 127  
DB 118 IESKIFNNWKDCQAVVYKRGITFKKFEQDLAENFISGMSADHYKLMNISKESEFSKYK 177  
QY 128 VSRDTFSYMGDPVYVYTTDCCSSNGRRKPRAGIGYVWGBGHLNVG-IRLPGRTNORAE 186  
DB 178 ISSNTM--YNSKMNYCCGSSFGNGTSSRAFGYVFGAPENISEPLLSAQTNRAE 235  
QY 187 IHAACKAIEQ--AKTQNTKLVLY---TDSMTTNGITNTWYQWKGKMGKTSAGKEVIN 239  
DB 236 IFAVSEALKKIWEKLTNEKKNYQIKTDSRYVTLANDRYVTYNTKKLGBLPNDLIVP 295  
QY 240 -----NKEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQ 283  
DB 296 LVGRFVYKVKYELNKECP-----KNNKGFOIEM--VKGHDDPGNEMADFLAKKASR 347





QY 188 HAACKAIBOAKTONINKLVLTYSMTFTINGITNNVOGKKNKMTSAGKEVINKEDFVL 247  
 Db 50 RAVIELALTLKEPCP--ITLVSDSQYKMGITKWLIFNMKKNWKSAGKPVKNQDLMLAL 107  
 QY 248 ERLTQGMIDQMHVPGHSGFTIGNEADRLAREGAK 282  
 Db 108 DESIQRHKINQWGVKGHAGHRENEICDELAKKGAE 142  
 RESULT 11  
 A83418  
 ribonuclease H PA1815 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: A83418  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lam,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: A83418  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-148 <STO>  
 A:Cross-references: GB:AE004608; GB:AE004091; NID:g9947797; PIDN:AA05204.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: rnhA; PA1815  
 C:Superfamily: ribonuclease H

Query Match 15.1%; Score 233.5; DB 2; Length 148;  
 Best Local Similarity 35.8%; Pred. No. 2e-13;  
 Matches 54; Conservative 22; Mismatches 56; Indels 19; Gaps 4;  
 QY 140 VVVVYTDGCCSSNGRRKPRAGIGVY-----WQPGHPLNVGIRLPGRQTNORAEIHAACK 192  
 Db 7 VVITTDGACKNPGKCGKALILYKGAERLNG--GEP-----DTNNKMEIQAALQ 56  
 QY 193 ALEQAKTONINKLVLTYSMTFTINGITNNVOGKKNKMTSAGKEVINKEDFVALERLQ 252  
 Db 57 ALAALKRSCPIRLI--TDSEYVVRGITEWLPNMKKRGKMTASKQPVKNADLQALDEQVA 114  
 QY 253 GMDIQMHVPGHSGFTIGNEADRLAREGAKQ 283  
 Db 115 RHQVEMQWVRGHTGDPGNERADQLANRGVAE 145

RESULT 12  
 F81807  
 ribonuclease H (EC 3.1.26.4) I NMA1817 [imported] - Neisseria meningitidis (strain Z2491)  
 C:Species: Neisseria meningitidis  
 C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C:Accession: F81807  
 R:Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
 ; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
 Nature 404, 502-506, 2000  
 A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A:Reference number: A81775; MUID:20222556; PMID:10761919  
 A:Accession: F81807  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-145 <PAR>  
 A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA885042.1; PID:g738045  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: rnhA; NMA1817  
 C:Superfamily: ribonuclease H  
 C:Keywords: hydrolase

Query Match 15.0%; Score 232.5; DB 2; Length 145;  
 Best Local Similarity 34.8%; Pred. No. 2.4e-13;

Matches 55; Conservative 22; Mismatches 56; Indels 25; Gaps 5;  
 QY 136 MDDPVVVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQ-----TNORA 185  
 Db 1 MNQTYVLTLDGACKN---PGAG---GWS-----VLMRFGSHKEKLPGEAQTNNNM 47  
 QY 186 EIHAAKCAIBOAKTONINKLVLTYSMTFTINGITNNVOGKKNKMTSAGKEVINKEDFV 245  
 Db 48 ELTAVIEGLKSLKRR--CTVIICITDSQYKNGEMWVHGMKKNKMTASKQPVKNDDLMK 105  
 QY 246 ALERLTQGMIDQMHVPGHSGFTIGNEADRLAREGAKQ 283  
 Db 106 ELDALVGRHQVSWTWVKGHAGHAENBRADDLANRGAQ 143

RESULT 13  
 H81061  
 ribonuclease HI NMB1618 [imported] - Neisseria meningitidis (strain MC58 serogroup B)  
 C:Species: Neisseria meningitidis  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: H81061  
 R:Petelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Frazer, C.M.; Moxon, E.R.; Rappuoli, R.; Ver  
 A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755; PMID:10720307  
 A:Accession: H81061  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-145 <TEF>  
 A:Cross-references: GB:AE002512; GB:AE002098; NID:g7226866; PIDN:AAFA1970.1; PID:g7226865  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1618  
 C:Superfamily: ribonuclease H

Query Match 14.8%; Score 229.5; DB 2; Length 145;  
 Best Local Similarity 34.2%; Pred. No. 4.5e-13;  
 Matches 54; Conservative 23; Mismatches 56; Indels 25; Gaps 5;  
 QY 136 MDDPVVVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQ-----TNORA 185  
 Db 1 MNQTYVLTLDGACKN---PGAG---GWS-----VLMRFGSHKEKLPGEAQTNNNM 47  
 QY 186 EIHAAKCAIBOAKTONINKLVLTYSMTFTINGITNNVOGKKNKMTSAGKEVINKEDFV 245  
 Db 48 ELTAVIEGLKSLKRR--CTVIICITDSQYKNGEMWVHGMKKNKMTASKQPVKNDDLMK 105  
 QY 246 ALERLTQGMIDQMHVPGHSGFTIGNEADRLAREGAKQ 283  
 Db 106 ELDALVGRHQVSWTWVKGHAGHAENBRADDLANRGAQ 143

RESULT 14  
 T33725  
 ribonuclease H (EC 3.1.26.4) - Zymomonas mobilis  
 C:Species: Zymomonas mobilis  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T33725  
 R:Lee, J.S.; Jin, S.J.; Kang, H.L.; Kang, H.S.  
 submitted to the EMBL Data Library, August 1998  
 A:Description: Sequence analysis of 67E10 cosmid clone of Zymomonas mobilis ZM4.  
 A:Reference number: Z21392  
 A:Accession: T33725  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-156 <LBE>  
 A:Cross-references: EMBL:AF086791; NID:g3820581; PID:g3089615; PIDN:AACT0364.1  
 C:Genetics:  
 A:Gene: rnh  
 C:Function:

A;Description: degrades the RNA of RNA-DNA hybrids specifically  
C;Superfamily: ribonuclease H  
C;Keywords: hydrolase

Query Match	14.5%	Score 224.5	DB 2	Length 156
Best Local Similarity	37.8%	Pred No. 1.4e-12		
Matches 56	Conservative 21	Mismatches 54	Indels 17	Gaps 6

RESULT 15  
AB2672

ribonuclease H [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont C-58000), *Agrobacterium tumefaciens*

C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C/Accession: AB2672  
R/Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, P.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCell  
; Kard, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
lter, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens*  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AB2672  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-146 <KUR>  
A:Cross-references: GB:AE006686; PIDN:AALL1792.1; PID:gl17739146; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atcu0776  
A:Map position: circular chromosome  
C:Superfamily: ribonuclease H  
  
Query Match 14.3% Score 220.5; DB 2; Length 146;  
Best Local Similarity 40.1%; Pred. No. 2.8e-12;  
Matches 59; Conservative 17; Mismatches 54; Indels 17; Gaps 5

## RESULT 16

H9/453

[illegible]

A:Accession number: A97359; MWID:21608551; PMID:11943194  
A:Reference: H97453  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1746 <R>  
A:Cross-references: GB:AEO07869; PIDDN:AAK6585.1; PID:g15155751; GSPDB:GN00169  
A:Genetics: C:117

A;gene: AGR\_C\_1417  
A;Map position: circular chromosome  
C;Superfamily: ribonuclease H

Query Match	14.3%	Score 220.5	DB 2	Length 146
Best Local Similarity	40.1%	Pred. No. 2,8e-12		
Matches 59	Conservative 17	Mismatches 54	Indels 17	Gaps 5

## RESULT 17

C87666  
ribonuclease HI [imported] - *Caulobacter crescentus*

C/Species: *Caulobacter crescentus*  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C/Accession: C87666  
R/Nielsen, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen, J.; Heidelberg, J.;  
B. Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolom  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A/Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87666  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-149 <STO>  
A;Cross-references: GB:AE005673, NID:G13425071, PIDN:AAK25327.1, GSPDB:GN00148  
C;Genetics:  
A;Gene: CC3365  
C;Superfamily: ribonuclease H

Query Match 14.2%; Score 219.5; DB 2; Length 149;  
Best Local Similarity 36.1%; Pred. No. 3.5e-12;  
Matches 56; Conservative 21; Mismatches 53; Indels 25; Gaps 6

## RESULT 18

AI0132

Call Enzymus ribonucleotidase H (BC 3.1.26.4) [imported] - Yersinia pestis (strain CO92)  
 C.Species: Yersinia pestis  
 C.Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
 C.Accession: A01032  
 E.Packhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
 Nature 413, 523-527, 2001  
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A'Accession: A10132  
 A'Reference number: AB0001; MUID:21470413; PMID:11586360  
 A>Status: preliminary  
 A'Molecule type: DNA  
 A'Residuals: 1-154 <KUR>  
 A'Cross-references: GB:AL590842; PDB:CA089924.1; PID:g15979149; GSPDB:GN00175  
 C'Genetics:  
 A'Gene: rnhA  
 C'Superfamily: ribonuclease H

Query Match 13.9%; Score 214.5; DB 2; Length 154;  
 Best Local Similarity 35.7%; Pred. No. 1e-11;  
 Matches 55; Conservative 16; Mismatches 54; Indels 29; Gaps 4;

QY 140 VVVYTDGCGSSNGRRKPRAGIGVWGPGRPLNVGIRLPGRQ-----TTQRAEI 187  
 DB 5 VEIIFDGSGLNPG-----GPG---GCGAILRKKQHEKTPSAGYVLTNNRML 49  
 QY 188 HAACKAIEQAKTQINKLVLYTDSMFTINGITNWVGKKGKTSAGKEVINKEDFVAL 247  
 DB 50 MAATVALLEALTSR--CEVTLSTDSQYVRGQITQWIMHWKKGKTPADRKPRVNVLDLQRL 107  
 QY 248 ERLTQGMIDQMMHVPGHSGFIGNBEADRLAREGA 281  
 DB 108 DLAIQSHITQWEMVKGHAGHPNERCDELARQGA 141

## RESULT 19

ribonuclease H VC2234 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
 C'Species: *Vibrio cholerae*  
 C'Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C'Accession: G82101  
 R.Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
 1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A>Title: DNA Sequence of both chromosomes of the *cholera* pathogen *Vibrio cholerae*.  
 A'Reference number: A82035; MUID:20406833; PMID:10952301  
 A'Accession: G82101  
 A>Status: preliminary  
 A'Molecule type: DNA  
 A'Residuals: 1-156 <HEI>  
 A'Cross-references: GB:AE004295; GB:AE003852; NID:99656789; PDB:AAF95378.1; GSPDB:GN001  
 A'Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C'Genetics:  
 A'Gene: VC2234  
 A'Map position: 1  
 C'Superfamily: ribonuclease H

Query Match 13.8%; Score 214; DB 2; Length 156;  
 Best Local Similarity 34.1%; Pred. No. 1.2e-11;  
 Matches 56; Conservative 21; Mismatches 57; Indels 30; Gaps 5;

QY 136 MGDFFVVYTDGCGSSNGRRKPRAGIGVWGP-----HPLNVGIRLPGRQTNQ 183  
 DB 1 MKQYVEIFDGSGLNPG-----GPGYGVIMRYKQVEKTLAAGYRL---TTNN 45  
 QY 184 RAETHAACAIEQAKTQINKLVLYTDSMFTINGITNWVGKKGKTSAGKEVINKED 243  
 DB 46 RHEMLAAVVALALKEP--CRVTLTDSQYVRGQITQWIMHWKKGKTPADRKPRVNVLDLQRL 103  
 QY 244 FVALERLTQGMIDQMMHVPGHSGFIGNBEADRLAREGAKO--SD 286  
 DB 104 WQALDKETARHGVEMRWVKGHAGHPNERCDELARQGALENPTD 147

## RESULT 20

NRECH

ribonuclease H (EC 3.1.26.4) - *Escherichia coli* (strain K-12)  
 C'Species: *Escherichia coli*  
 C'Date: 18-Apr-1984 #sequence\_revision 18-Apr-1984 #text\_change 01-Mar-2002  
 C'Accession: A92401; A93979; S13170; B24257; I54647; H64745; A00793  
 R.Kanaya, S.; Crouch, R.U.  
 J. Biol. Chem. 258, 1276-1281, 1983  
 A>Title: DNA sequence of the gene coding for *Escherichia coli* ribonuclease H.  
 A'Reference number: A92401; MUID:83108846; PMID:6296074  
 A'Accession: A92401  
 A'Molecule type: DNA  
 A'Residuals: 1-155 <KAN>  
 A'Cross-references: GB:K00552; GB:U01676; GB:V00337; NID:g147676; PID:g147677  
 R.Maki, H.; Horuchi, T.; Sekiguchi, M.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 7137-7141, 1983  
 A>Title: Structure and expression of the dnaQ mutator and the Rnae H genes of *Escherichia coli*.  
 A'Reference number: A93979; MUID:84070781; PMID:6316347  
 A'Accession: A93979  
 A'Molecule type: DNA  
 A'Residuals: 1-155 <MAK>  
 A'Cross-references: GB:K00985; GB:M30201; NID:g147678; PDB:AAA24565.1; PID:g147680  
 R.Kanaya, S.; Kimura, S.; Karsuda, C.; Ikehara, M.  
 Biochem. J. 271, 59-66, 1990  
 A>Title: Role of cysteine residues in ribonuclease H from *Escherichia coli*.  
 A'Reference number: S13170; MUID:91024947; PMID:2171503  
 A'Accession: S13170  
 A'Molecule type: DNA  
 A'Residuals: 1-155 <KA2>  
 R.Cox, E.C.; Horner, D.L.  
 J. Mol. Biol. 190, 113-117, 1986  
 A>Title: DNA sequence and coding properties of mutD(dnaQ) a dominant *Escherichia coli* mut  
 A'Reference number: A24257; MUID:87060973; PMID:3023634  
 A'Accession: B24257  
 A'Molecule type: DNA  
 A'Residuals: 1-155 <COX>  
 A'Cross-references: GB:X04027; NID:942061; PDB:CAA27660.1; PID:g42062  
 J.Kanaya, S.; Crouch, R.U.  
 J. Bacteriol. 154, 1021-1026, 1983  
 A>Title: Low levels of Rnae H activity in *Escherichia coli* FB2 rnh result from a single  
 A'Reference number: I54847; MUID:83185598; PMID:6302075  
 A'Accession: I54847  
 A'Molecule type: DNA  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A'Residuals: 1-155 <RES>  
 A'Cross-references: EMBL:V00337; NID:942776; PDB:CAA23620.1; PID:g42777  
 R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A>Title: The complete genome sequence of *Escherichia coli* K-12.  
 A'Reference number: A64720; MUID:97426617; PMID:9278503  
 A'Accession: H64745  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A'Molecule type: DNA  
 A'Residuals: 1-155 <BLAT>  
 A'Cross-references: GB:AE000130; GB:U00096; NID:g1786402; PDB:AACT3319.1; PID:g1786408;  
 A'Experimental source: strain K-12, substrain MG1655  
 C'Genetics:  
 A'Gene: rnhA, rnh  
 A'Map position: 5 min  
 C'Function:  
 A>Description: an endonuclease that degrades the RNA of RNA-DNA hybrids specifically  
 C'Superfamily: ribonuclease H  
 C'keyword: hydrolase

Query Match 13.5%; Score 208.5; DB 1; Length 155;  
 Best Local Similarity 36.6%; Pred. No. 3.5e-11;  
 Matches 52; Conservative 16; Mismatches 69; Indels 5; Gaps 2;

QY 140 VVVYTDGCGSSNGRRKPRAGIGVWGPGRPLNVGIRLPGRQNRORAEIHAACKAIEQAKT 199  
 DB 5 VEIIFDGSGLNPGCGGAILIRYRGREKRTTSAGT---RTNNRMLAAVVALBALKE 61  
 QY 200 QNINKLVLYTDSMFTINGITNWVGKKGKTSAGKEVINKEDFVALERLTQGMIDQMM 259  
 DB 100 QNINKLVLYTDSMFTINGITNWVGKKGKTSAGKEVINKEDFVALERLTQGMIDQMM 259



QY 200 QNINKLVLTDSMTFTINGTNVQGMKKGWTSAGKEVINKEDFVALERLTQGMIDIOM 259  
DB 62 H--CEVTLSSTDSQYRQGTOMIHNMKKKGWTKAEKPKVNDLWKRLDPAALGQHQRKV 119  
QY 260 HVPGHSGFTIGNEADRLAREGA 281  
DB 120 WYKGHAGHPENRRCDELAARAAA 141

## RESULT 25

F82591

ribonuclease H XF2158 [imported] - Xylella fastidiosa (strain 9asc)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000

C:Accession: F82591

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82591

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 &lt;SIM&gt;

A:Cross-references: GB:AE004030; GB:AE003849; NID:g9107292; PIDN:AAf84957.1; GSPDB:GN001

A:Experimental source: strain 9asc

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.B.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Vertovaki-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2158

C:Superfamily: ribonuclease H

Query Match 12.6%; Score 194.5; DB 2; length 153;

Best Local Similarity 35.2%; Pred. No. 6e-10;

Matches 51; Conservative 17; Mismatches 54; Indels 23; Gaps 4;

QY 143 YTDGCCSSNGRRKPRAGIGVWGP-----HPINVGIRLPGRQ---TNGRAEIHACKA 193  
DB 10 YTDGSCLGNP-----GPGGMAVLLRYKNNKEKLVGGLDTNNRMELMAAIA 57  
QY 194 IEQAKTQNIKLVLTDSMTFTINGTNVQGMKKGWTSAGKEVINKEDFVALERLTQGMIDIOM 253  
DB 58 LE--RLSEPCQIKLHTDSQYRQGTOMIHNMKKKGWTKAEKPKVNDLWKRLDPAALGQHQRKV 119  
QY 254 MDIQMHVPGHSGFTIGNEADRLAREGA 281  
DB 116 HMYEWCMVKAHNGDSNDERVDYLAR 140

Search completed: October 7, 2004, 08:09:10  
Job time : 45 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:08:00 ; Search time 127 Seconds  
(without alignments)  
724.662 Million cell updates/sec

Title: US-10-054-313-1

Perfect score: 1546  
Sequence: 1 MSWLLFLAHRLAALALPCRR.....FIGNBEADRLAREGKQSD 286

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1546	100.0	286	9	US-09-861-205-1 Sequence 1, Appli
2	1546	100.0	286	13	US-10-054-313-1 Sequence 1, Appli
3	1546	100.0	286	14	US-10-358-439-1 Sequence 1, Appli
4	1522	98.4	286	12	US-10-262-511-86 Sequence 86, Appli
5	1518	98.2	286	14	US-10-358-439-6 Sequence 6, Appli
6	1420	91.8	286	12	US-09-781-712B-6 Sequence 6, Appli
7	1407	91.0	286	12	US-09-781-712B-7 Sequence 7, Appli
8	1403	90.8	286	12	US-09-781-712B-8 Sequence 8, Appli
9	1399	90.5	286	12	US-09-781-712B-9 Sequence 9, Appli
10	1086.5	70.3	285	12	US-09-781-712B-11 Sequence 11, Appli
11	1045.5	67.6	203	12	US-10-262-511-78 Sequence 78, Appli
12	1045.5	67.6	210	12	US-10-262-511-80 Sequence 80, Appli
13	979.5	63.4	195	12	US-10-262-511-82 Sequence 82, Appli
14	959	62.0	216	9	US-09-861-205-5 Sequence 5, Appli
15	959	62.0	216	13	US-10-054-313-5 Sequence 5, Appli

16	959	62.0	216	14	US-10-358-439-5 Sequence 5, Appli
17	882	57.1	293	14	US-10-358-439-2 Sequence 2, Appli
18	880	56.9	293	13	US-09-861-205-2 Sequence 2, Appli
19	880	56.9	293	13	US-10-054-313-2 Sequence 2, Appli
20	792	51.2	152	12	US-10-262-511-84 Sequence 84, Appli
21	440.5	28.5	369	15	US-10-369-493-5383 Sequence 5383, Ap
22	424.5	27.5	189	9	US-09-992-738-1 Sequence 1, Appli
23	335	21.7	300	15	US-10-369-493-13271 Sequence 13271, A
24	310	20.1	275	15	US-10-369-493-1170 Sequence 1170, Ap
25	288	18.6	325	15	US-10-369-493-2540 Sequence 2540, Ap
26	249.5	16.1	348	9	US-09-861-205-3 Sequence 3, Appli
27	249.5	16.1	348	13	US-10-054-313-3 Sequence 3, Appli
28	249.5	16.1	348	14	US-10-358-439-3 Sequence 3, Appli
29	249.5	16.1	348	15	US-10-369-493-1914 Sequence 1914, Ap
30	235.5	15.3	154	14	US-10-260-877-72 Sequence 72, Appli
31	235.5	15.3	266	10	US-09-975-719-157 Sequence 157, App
32	208.5	13.5	155	9	US-09-861-205-4 Sequence 4, Appli
33	208.5	13.5	155	13	US-10-054-313-4 Sequence 4, Appli
34	208.5	13.5	155	14	US-10-358-439-4 Sequence 4, Appli
35	176.5	11.4	236	14	US-10-156-761-8677 Sequence 8677, Ap
36	158	10.2	223	15	US-10-369-493-3018 Sequence 3018, Ap
37	152	9.8	1199	16	US-10-677-558-2 Sequence 2, Appli
38	137	8.9	716	9	US-09-845-157-2 Sequence 2, Appli
39	133.5	8.6	1224	16	US-10-437-963-175172 Sequence 175172, A
40	133.5	8.6	1290	16	US-10-437-963-175572 Sequence 175572, A
41	133	8.6	1707	16	US-10-437-963-149578 Sequence 149578, A
42	129	8.3	1547	16	US-10-437-963-138161 Sequence 138161, A
43	124.5	8.1	1018	14	US-10-369-294-46 Sequence 46, Appli
44	123	8.0	1279	16	US-10-437-963-138178 Sequence 138178, A
45	122.5	7.9	1258	16	US-10-437-963-162179 Sequence 162179, A
46	122.5	7.9	1678	16	US-10-437-963-138217 Sequence 138217, A
47	120	7.8	1053	16	US-10-437-963-175580 Sequence 175580, A
48	119	7.7	1383	16	US-10-437-963-138172 Sequence 138172, A
49	117.5	7.6	1047	16	US-10-437-963-138177 Sequence 138177, A
50	116.5	7.5	711	16	US-10-437-963-138135 Sequence 138135, A
51	116.5	7.5	1016	14	US-10-364-360-24 Sequence 24, Appli
52	116.5	7.5	1898	16	US-10-437-963-138220 Sequence 138220, A
53	116	7.5	549	16	US-10-437-963-179210 Sequence 179210, A
54	114.5	7.4	1495	16	US-10-437-963-138039 Sequence 138039, A
55	114.5	7.4	1495	16	US-10-437-963-138063 Sequence 138063, A
56	114	7.4	321	14	US-10-029-386-31960 Sequence 31960, A
57	113.5	7.3	1865	16	US-10-437-963-138159 Sequence 138159, A
58	112.5	7.3	1322	16	US-10-437-963-138139 Sequence 138139, A
59	112	7.2	1201	16	US-10-437-963-105133 Sequence 105133, A
60	111.5	7.2	143	10	US-09-882-227-54 Sequence 54, Appli
61	111.5	7.2	228	9	US-09-864-761-36315 Sequence 36315, A
62	111.5	7.1	1573	16	US-10-437-963-192716 Sequence 192716, A
63	110.5	7.1	1828	16	US-10-437-963-152118 Sequence 152118, A
64	109.5	7.1	1449	16	US-10-437-963-105135 Sequence 105135, A
65	109	7.1	1442	16	US-10-437-963-150849 Sequence 150849, A
66	109	7.1	1452	16	US-10-437-963-138097 Sequence 138097, A
67	109	7.1	1933	16	US-10-437-963-138059 Sequence 148359, A
68	108	7.0	704	16	US-10-437-963-149914 Sequence 149914, A
69	108	7.0	1265	16	US-10-437-963-138094 Sequence 138094, A
70	108	7.0	1930	16	US-10-437-963-138142 Sequence 138142, A
71	108	7.0	2012	16	US-10-437-963-164042 Sequence 164042, A
72	107.5	7.0	1014	14	US-10-301-651A-6 Sequence 6, Appli
73	107	6.9	995	12	US-10-296-734-2 Sequence 2, Appli
74	107	6.9	995	12	US-10-296-734-1470 Sequence 1470, Ap
75	107	6.9	1130	16	US-10-437-963-144806 Sequence 144806, A
76	107	6.9	1150	9	US-09-946-239-9 Sequence 9, Appli
77	107	6.9	1902	16	US-10-437-963-138141 Sequence 138141, A
78	107	6.9	1584	16	US-10-437-963-138095 Sequence 138095, A
79	107	6.9	1919	16	US-10-437-963-133285 Sequence 133285, A
80	106.5	6.9	999	15	US-10-346-000A-3 Sequence 3, Appli
81	106.5	6.9	1040	16	US-10-437-963-162180 Sequence 162180, A
82	106	6.9	1533	9	US-10-437-963-146756 Sequence 146756, A
83	105.5	6.8	263	16	US-09-864-761-33879 Sequence 33879, A
84	105.5	6.8	1790	16	US-10-437-963-138166 Sequence 138166, A
85	105	6.8	332	12	US-10-437-963-118142 Sequence 118142, A
86	104.5	6.8	143	16	US-10-335-977-6991 Sequence 6991, Ap
87	104.5	6.8	163	12	US-10-335-977-6992 Sequence 6992, Ap
88	103.5	6.7	1940	16	US-10-437-963-149581 Sequence 149581, A

89 103 6.7 998 16 US-10-332-413-4 Sequence 4, Appli  
90 103 6.7 1734 16 US-10-437-963-168292 Sequence 168292,  
91 102 6.6 1857 16 US-10-437-963-152007 Sequence 152007,  
92 101.5 6.6 1306 16 US-10-437-963-146656 Sequence 146656,  
93 101 6.5 940 16 US-10-437-963-162186 Sequence 162186,  
94 101 6.5 1082 16 US-10-437-963-162190 Sequence 162190,  
95 101 6.5 1424 16 US-10-437-963-160683 Sequence 160683,  
96 101 6.5 1733 16 US-10-437-963-193694 Sequence 193694,  
97 101 6.5 1752 16 US-10-437-963-168295 Sequence 168295,  
98 100.5 6.5 1217 16 US-10-437-963-133329 Sequence 133329,  
99 100 6.5 560 12 US-10-399-920-2 Sequence 2, Appli  
100 100 6.5 560 12 US-09-725-652-1 Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-861-205-1  
; Sequence 1, Application US/09861205  
; Patent No. US20020076712A1  
; GENERAL INFORMATION:  
; APPLICANT: Crooke, Stanley T.  
; APPLICANT: Lima, Walter F.  
; APPLICANT: Wu, Hongliang  
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
; FILE REFERENCE: ISPH-0333  
; CURRENT APPLICATION NUMBER: US/09/861,205  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/203,716  
; PRIOR FILING DATE: 1998-12-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-861-205-1

Query Match 100.0%; Score 1546; DB 9; Length 286;  
Best Local Similarity 100.0%; Pred. No. 8.8e-152;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLARVALAALPCRRGSGFGMFYAVRRGRKTGVFLTNNECRAQVDRPPAARFKK 60  
DB 1 MSWLLFLARVALAALPCRRGSGFGMFYAVRRGRKTGVFLTNNECRAQVDRPPAARFKK 60  
QY 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAQPYAKHMKP 120  
DB 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAQPYAKHMKP 120  
QY 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
DB 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
QY 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTINGITNWVGKKNKGTSGAGKEVIN 240  
DB 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTINGITNWVGKKNKGTSGAGKEVIN 240  
QY 241 KEDFVALERLTQGMIDQMHVPGHSGFIGNEEADRLAREGAKOSED 286  
DB 241 KEDFVALERLTQGMIDQMHVPGHSGFIGNEEADRLAREGAKOSED 286

RESULT 2  
US-10-054-313-1  
; Sequence 1, Application US/10054313  
; Publication No. US20020110892A1  
; GENERAL INFORMATION:  
; APPLICANT: Crooke, Stanley T.  
; APPLICANT: Lima, Walter F.  
; APPLICANT: Wu, Hongliang  
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof

; FILE REFERENCE: ISPH-0333  
; CURRENT APPLICATION NUMBER: US/10/054,313  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/203,716  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-054-313-1

Query Match 100.0%; Score 1546; DB 13; Length 286;  
Best Local Similarity 100.0%; Pred. No. 8.8e-152;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLARVALAALPCRRGSGFGMFYAVRRGRKTGVFLTNNECRAQVDRPPAARFKK 60  
DB 1 MSWLLFLARVALAALPCRRGSGFGMFYAVRRGRKTGVFLTNNECRAQVDRPPAARFKK 60  
QY 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAQPYAKHMKP 120  
DB 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAQPYAKHMKP 120  
QY 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
DB 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
QY 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTINGITNWVGKKNKGTSGAGKEVIN 240  
DB 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTINGITNWVGKKNKGTSGAGKEVIN 240  
QY 241 KEDFVALERLTQGMIDQMHVPGHSGFIGNEEADRLAREGAKOSED 286  
DB 241 KEDFVALERLTQGMIDQMHVPGHSGFIGNEEADRLAREGAKOSED 286

RESULT 3  
US-10-358-439-1  
; Sequence 1, Application US/10358439  
; Publication No. US20030144486A1  
; GENERAL INFORMATION:  
; APPLICANT: Crooke, Stanley T.  
; APPLICANT: Lima, Walter F.  
; APPLICANT: Wu, Hongliang  
; TITLE OF INVENTION: Human RNase H and Compositions And Uses Thereof  
; FILE REFERENCE: ISPH0725  
; CURRENT APPLICATION NUMBER: US/10/358,439  
; CURRENT FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/861,205  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/584,254  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/343,809  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/203,716  
; PRIOR FILING DATE: 1998-12-02  
; PRIOR APPLICATION NUMBER: 60/067,458  
; PRIOR FILING DATE: 1997-12-04  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-358-439-1

Query Match 100.0%; Score 1546; DB 14; Length 286;  
Best Local Similarity 100.0%; Pred. No. 8.8e-152;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLARVALAALPCRRGSGFGMFYAVRRGRKTGVFLTNNECRAQVDRPPAARFKK 60



```

Db 1 MSWLF.LAHRYVALA.LP.CRRGSRGFMFYAVRGRKTVFLTWNCRAQVDRFPAPARFKK 60
Qy 61 FATEDEAMAFVRRKASPEVSEGEHNOQESSEAKGKRLREPLDDGDHESAPYAKHMKP 120
Db 61 FATEDEAMAFVRRKASPEVSEGEHNOQESSEAKGKRLREPLDDGDHESAPYAKHMKP 120
Qy 121 SVEPAPVSRDPTFSYMGDPVVVYTTDCCSSNGRRPRRAGIGYWGPGHPLNVGIRLPGRQ 180
Db 121 SVEPAPVSRDPTFSYMGDPVVVYTTDCCSSNGRRPRRAGIGYWGPGHPLNVGIRLPGRQ 180
Qy 181 TNORAEIHAAKCAIQAQATQINIKLVLYTDSMFTINGITNNVQKKGKMTSACKEVIN 240
Db 181 TNORAEIHAAKCAIQAQATQINIKLVLYTDSMFTINGITNNVQKKGKMTSACKEVIN 240
Qy 241 KEDFVALERLTQGMIDIQMWHVPGHSGFTIGNEADRLAREGAKOSBD 286
Db 241 KEDFVALERLTQGMIDIQMWHVPGHSGFTIGNEADRLAREGAKOSBD 286

```

## RESULT 4

```

US-10-262-511-86
; Sequence 86, Application US/10262511
; Publication No. US2004003823A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Miller, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jinfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Saeha)
; APPLICANT: Patumraj, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shilmit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malysankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zetnusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 2102-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260

```

```

; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Cureseq1ist version 0.1
; SEQ ID NO 86
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-86

```

```

Query Match 98.4%; Score 1522; DB 12; Length 286;
Best Local Similarity 98.3%; Pred. No. 2.8e-143;
Matches 281; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 MSWLF.LAHRYVALA.LP.CRRGSRGFMFYAVRGRKTVFLTWNCRAQVDRFPAPARFKK 60
Db 1 MSWLF.LAHRYVALA.LP.CRRGSRGFMFYAVRGRKTVFLTWNCRAQVDRFPAPARFKK 60
Qy 61 FATEDEAMAFVRRKASPEVSEGEHNOQESSEAKGKRLREPLDDGDHESAPYAKHMKP 120
Db 61 FATEDEAMAFVRRKASPEVSEGEHNOQESSEAKGKRLREPLDDGDHESAPYAKHMKP 120
Qy 121 SVEPAPVSRDPTFSYMGDPVVVYTTDCCSSNGRRPRRAGIGYWGPGHPLNVGIRLPGRQ 180
Db 121 SVEPAPVSRDPTFSYMGDPVVVYTTDCCSSNGRRPRRAGIGYWGPGHPLNVGIRLPGRQ 180
Qy 181 TNORAEIHAAKCAIQAQATQINIKLVLYTDSMFTINGITNNVQKKGKMTSACKEVIN 240
Db 181 TNORAEIHAAKCAIQAQATQINIKLVLYTDSMFTINGITNNVQKKGKMTSACKEVIN 240
Qy 241 KEDFVALERLTQGMIDIQMWHVPGHSGFTIGNEADRLAREGAKOSBD 286
Db 241 KEDFVALERLTQGMIDIQMWHVPGHSGFTIGNEADRLAREGAKOSBD 286

```

## RESULT 5

```

US-10-358-439-6
; Sequence 6, Application US/10358439
; Publication No. US2003014496A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H and Compositions And Uses Thereof
; FILE REFERENCE: ISPH0725
; CURRENT APPLICATION NUMBER: US/10/358,439
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/861,205
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/684,254
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/343,809
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 60/067,458
; PRIOR FILING DATE: 1997-12-04
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (95)..(96)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature

```

LOCATION: (112)..(112)  
 OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (155)..(155)  
 OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 US-10-358-439-6

Query Match 98.2%; Score 1518; DB 14; Length 286;  
 Best Local Similarity 98.6%; Pred. No. 7.2e-149;  
 Matches 282; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 DB 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 QY 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120  
 DB 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120  
 QY 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 QY 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240  
 DB 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240  
 QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286  
 DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286

RESULT 6  
 US-09-781-712B-6  
 Sequence 6, Application US/09781712B  
 Publication No. US20040180433A1

GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T  
 APPLICANT: Lima, Walter  
 TITLE OF INVENTION: Method of Using Mammalian RNase H and Compositions Thereof  
 FILE REFERENCE: ISPH-0520  
 CURRENT APPLICATION NUMBER: US/09/781,712B  
 PRIOR FILING DATE: 2001-02-12  
 PRIOR APPLICATION NUMBER: US 60/067,458  
 PRIOR FILING DATE: 1997-12-04  
 PRIOR APPLICATION NUMBER: US 09/203,716  
 PRIOR FILING DATE: 1998-12-02  
 PRIOR APPLICATION NUMBER: US 09/343,809  
 PRIOR FILING DATE: 1999-06-30  
 PRIOR APPLICATION NUMBER: US 09/684,254  
 PRIOR FILING DATE: 2000-10-06  
 NUMBER OF SEQ ID NOS: 39  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 6  
 LENGTH: 286  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-781-712B-6

Query Match 91.8%; Score 1420; DB 12; Length 286;  
 Best Local Similarity 93.7%; Pred. No. 1.1e-138;  
 Matches 268; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 DB 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 QY 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120  
 DB 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120

QY 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 QY 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240  
 DB 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240

QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286  
 DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286

RESULT 7  
 US-09-781-712B-7

Sequence 7, Application US/09781712B  
 Publication No. US20040180433A1  
 GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T  
 APPLICANT: Lima, Walter  
 TITLE OF INVENTION: Method of Using Mammalian RNase H and Compositions Thereof  
 FILE REFERENCE: ISPH-0520  
 CURRENT APPLICATION NUMBER: US/09/781,712B  
 CURRENT FILING DATE: 2001-02-12  
 PRIOR APPLICATION NUMBER: US 60/067,458  
 PRIOR FILING DATE: 1997-12-04  
 PRIOR APPLICATION NUMBER: US 09/203,716  
 PRIOR FILING DATE: 1998-12-02  
 PRIOR APPLICATION NUMBER: US 09/343,809  
 PRIOR FILING DATE: 1999-06-30  
 PRIOR APPLICATION NUMBER: US 09/684,254  
 PRIOR FILING DATE: 2000-10-06  
 NUMBER OF SEQ ID NOS: 39  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 7  
 LENGTH: 286  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-781-712B-7

Query Match 91.0%; Score 1407; DB 12; Length 286;  
 Best Local Similarity 92.7%; Pred. No. 2.4e-137;  
 Matches 265; Conservative 2; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 DB 1 MSWLLFLAHVVALAALPCRRGSRGFGMFYAVRRGRKTGVLTWNECAQVDRPPAARFKK 60  
 QY 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120  
 DB 61 FATEDENAMAFVRKSASPEVSEGHENHGOSESAKPGKRLREPLDGDGHESAOFPYAKMKP 120  
 QY 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPPVSRDTSYMGDFVYVYTDGCCSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 QY 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240  
 DB 181 TNORAEIHAACKIAEQAKTONINKLVLYTDSMTFTINGITWVGKKNKWTSAKEVIN 240  
 QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286  
 DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286

RESULT 8  
 US-09-781-712B-8

Sequence 8, Application US/09781712B  
 Publication No. US20040180433A1  
 GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T  
 APPLICANT: Lima, Walter

APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Methods of Using Mammalian Rhase H and Compositions Thereof  
FILE REFERENCE: ISPH-0520  
CURRENT APPLICATION NUMBER: US 09/781,712B  
CURRENT FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/067,458  
PRIOR FILING DATE: 1997-12-04  
PRIOR APPLICATION NUMBER: US 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: US 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 09/684,254  
PRIOR FILING DATE: 2000-10-06  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-781-712B-8

Query Match 90.8%; Score 1403; DB 12; Length 286;  
Best Local Similarity 92.3%; Pred. No. 6.3e-137;  
Matches 264; Conservative 2; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRRGRKGVFLTWNECRQAVDRFPAPAFKK 60  
DB 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRRGRKGVFLTWNECRQAVDRFPAPAFLL 60  
QY 61 FATEDENAFVFKASPEVSEBGENHQESAEKPKRLREPLDDGHSAPYAKHMP 120  
DB 61 FATEDENAFVFKASPEVSEBGENHQESAEKPKRLREPLDDGHSAPYALHMP 120  
QY 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
DB 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
QY 181 TNQRAEIHAAKCAIEOAKTONINKLVLYTDSMFTINGITNNVQKKGWKTSAKEVIN 240  
DB 181 TNQRAEIHAAKCAIEOAKTONINKLVLYTDSMFTINGITNNVQKKGWKTSAKEVIN 240  
QY 241 KEDFVALERLTQGMNDIQMWHVPGHSGFTIGNEBADRAREGAKQSED 286  
DB 241 KEDFVALERLTQGMNDIQMWHVPGHSGFTIGNEBADRAREGAKQSED 286

RESULT 9  
US-09-781-712B-9  
Sequence 9, Application US/09781712B  
Publication No. US20040180433A1  
GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Methods of Using Mammalian Rhase H and Compositions Thereof  
FILE REFERENCE: ISPH-0520  
CURRENT APPLICATION NUMBER: US/09/781,712B  
CURRENT FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/067,458  
PRIOR FILING DATE: 1997-12-04  
PRIOR APPLICATION NUMBER: US 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: US 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 09/684,254  
PRIOR FILING DATE: 2000-10-06  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-781-712B-9

Query Match 90.5%; Score 1399; DB 12; Length 286;  
Best Local Similarity 92.0%; Pred. No. 1.7e-136;  
Matches 263; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRRGRKGVFLTWNECRQAVDRFPAPAFKK 60  
DB 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRRGRKGVFLTWNECRQAVDRFPAPAFLL 60  
QY 61 FATEDENAFVFKASPEVSEBGENHQESAEKPKRLREPLDDGHSAPYAKHMP 120  
DB 61 FATEDENAFVFKASPEVSEBGENHQESAEKPKRLREPLDDGHSAPYALHMP 120  
QY 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
DB 121 SVEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
QY 181 TNQRAEIHAAKCAIEOAKTONINKLVLYTDSMFTINGITNNVQKKGWKTSAKEVIN 240  
DB 181 TNQRAEIHAAKCAIEOAKTONINKLVLYTDSMFTINGITNNVQKKGWKTSAKEVIN 240  
QY 241 KEDFVALERLTQGMNDIQMWHVPGHSGFTIGNEBADRAREGAKQSED 286  
DB 241 KEDFVALERLTQGMNDIQMWHVPGHSGFTIGNEBADRAREGAKQSED 286

RESULT 10  
US-09-781-712B-11  
Sequence 11, Application US/09781712B  
Publication No. US20040180433A1  
GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Methods of Using Mammalian Rhase H and Compositions Thereof  
FILE REFERENCE: ISPH-0520  
CURRENT APPLICATION NUMBER: US/09/781,712B  
CURRENT FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/067,458  
PRIOR FILING DATE: 1997-12-04  
PRIOR APPLICATION NUMBER: US 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: US 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 09/684,254  
PRIOR FILING DATE: 2000-10-06  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 11  
LENGTH: 285  
TYPE: PRT  
ORGANISM: Mus sp.  
US-09-781-712B-11  
Query Match 70.3%; Score 1086.5; DB 12; Length 285;  
Best Local Similarity 72.7%; Pred. No. 4.8e-104;  
Matches 208; Conservative 22; Mismatches 55; Indels 1; Gaps 1;

Qy 241 KEDFVALERLTQGMIDQMHVPGHSGFTGNEADRLAREGAKQSED 286  
Db 240 LQDFMEDELTOGMIDQMHVPGHSGFTGNEADRLAREGAKQSED 285

## RESULT 11

US-10-262-511-78  
; Sequence 78, Application US/10262511  
; Publication No. US20040038223A1  
GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ji, Jinfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zernusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkete, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CursSeqdist version 0.1  
; SEQ ID NO: 78  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-262-511-78

Query Match 67.6%; Score 1045.5; DB 12; Length 203;  
Best Local Similarity 70.3%; Pred. No. 5.4e-100; Indels 83; Gaps 1;  
Matches 201; Conservative 1; Mismatches 1;

Qy 1 MSWLFLLARVVAALPCRRGSRGFGMFYAVRRGRTGVLFTNNEGRAQVDRRPAARFKK 60  
Db 1 MSWFLFLARVVAALPCRRGSRGFGMFYAVRRGRTGVLFTNNEC----- 46  
Qy 61 FATEDAMAFVRKASPEVSEGHENQGESEAKPKRLREPLDGDGHSAPYAKMKP 120  
Db 47 ----- 46  
Qy 121 SVEPAPVSRDPTSRWGDFFVVTYTDCCSSNGRRKRPAGIGYTWGPHPLNIGIRLRPGRQ 180  
Db 47 -----RDFTSYGDFVVTYTDCCSSNGRRRPAGIGYTWGPHPLNIGIRLRPGRQ 97  
Qy 181 TNGRAETHAKCAIEQAKTONINKLVLYTDSMTFTGTTNWVGKNGKTKTSAGKEVIN 240  
Db 98 TNGRAETHAKCAIEQAKTONINKLVLYTDSMTFTGTTNWVGKNGKTKTSAGKEVIN 157  
Qy 241 KEDFVALERLTQGMIDQMHVPGHSGFTGNEADRLAREGAKQSED 286  
Db 158 KEDFVALERLTQGMIDQMHVPGHSGFTGNEADRLAREGAKQSED 203

## RESULT 12

US-10-262-511-80  
; Sequence 80, Application US/10262511  
; Publication No. US20040038223A1  
GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ji, Jinfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zernusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkete, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09

```

; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining prior application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curoseqdist version 0.1
; SEQ ID NO 80
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-511-80

Query Match      67.6%; Score 1045.5; DB 12; Length 210;
Best Local Similarity 70.3%; Pred. No. 5.7e-100;
Matches 201; Conservative 1; Mismatches 1; Indels 83; Gaps 1;

QY 1 MSWLLFLARVALAALPCRRGSRGFMFYAVRGRKTGVFLTNWECRAQVDFPAPAFKK 60
DB 5 MSWPLFLARVALAALPCRRGSRGFMFYAVRGRKTGVFLTNWECRAQVDFPAPAFKK 50
QY 61 FATEDAMAFVRKKSAPVSEGEHNOGDESEAKPKRLREPLDGDGHSADQYAKHMKP 120
DB 51 -----
QY 121 SVEPAPVSRDPTFSYMGDFVYVYTDGCCSSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180
DB 51 -----RDTFSYMGDFVYVYTDGCCSSNGRRRPRAGIGYVWGPHPLNVGIRLPGRQ 101
QY 181 TNQRAEIIHAACKAIEQAKTONINKLVLYTDSMFTINGITNWVQMKKNGKTSAGKEVIN 240
DB 102 TNQRAEIIHAACKAIEQAKTONINKLVLYTDSMFTINGITNWVQMKKNGKTSAGKEVIN 161
QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286
DB 162 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRLAREGAKQSED 207

RESULT 13
; US-10-262-511-82
; Sequence 82, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Paturajan, Weera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zehusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Gatterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
```

```

; APPLICANT: Shinkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining prior application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curoseqdist version 0.1
; SEQ ID NO 82
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-262-511-82

Query Match      63.4%; Score 979.5; DB 12; Length 195;
Best Local Similarity 68.9%; Pred. No. 3.7e-93;
Matches 188; Conservative 2; Mismatches 0; Indels 83; Gaps 1;

QY 14 AALPCRRGSRGFMFYAVRGRKTGVFLTNWECRAQVDFPAPAFKKFATEDAMAFVRK 73
DB 3 AALPCRRGSRGFMFYAVRGRKTGVFLTNWECRAQVDFPAPAFKKFATEDAMAFVRK 35
QY 74 SASPEVSEGEHNOGDESEAKPKRLREPLDGDGHSADQYAKHMKPSVEPAPVSRDPT 133
DB 36 -----RDTF 39
QY 134 SYMGDFVYVYTDGCCSSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQTNQRAEIIHAACKA 193
DB 40 SYMGDFVYVYTDGCCSSNGRRRPRAGIGYVWGPHPLNVGIRLPGRQTNQRAEIIHAACKA 99
QY 194 ISOAKTONINKLVLYTDSMFTINGITNWVQMKKNGKTSAGKEVINKEDFVALERLTQ 253
DB 100 ISOAKTONINKLVLYTDSMFTINGITNWVQMKKNGKTSAGKEVINKEDFVALERLTQ 159
QY 254 MDIQMHVPGHSGFTIGNEADRLAREGAKQSED 286
DB 160 MDIQMHVPGHSGFTIGNEADRLAREGAKQSED 192

RESULT 14
; US-09-861-205-5
; Sequence 5, Application US/09861205
; Patent No. US20020076712A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongliang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0333
```

```

; CURRENT APPLICATION NUMBER: US/09/861,205
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-861-205-5

```

```

Query Match      62.0%; Score 959; DB 9; Length 216;
Best Local Similarity 70.2%; Pred. No. 5,7e-91;
Matches 177; Conservative 20; Mismatches 19; Indels 36; Gaps 1;

```

```

Qy 21 GSRGFGFYAVRRKRTGVFLTNWECRAQVDRPPAARFKKFPATEDEMAVFRKASPEVS 80
Db 1 GICGLGMFYAVRRGRRRPVEFLSWSECKAQVDRPPAARFKKFPATEDEMAVFRSSSPDGS 60
Qy 81 EGHENQGOSEAKPGKRLREPLDGDGHSQAQVPAKMKPSVEBAPPVSRDTFSYMGDFV 140
Db 61 KGOESAHQKSOAKTSRPREPL-----V 84
Qy 141 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAIEOAKTQ 200
Db 85 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAVMOAKAQ 144
Qy 201 NINKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 260
Db 145 NISKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 204
Qy 261 VPGHSGFTGNNE 272
Db 205 IPGHSGFVGNEE 216

```

```

RESULT 15
US-10-054-313-5
; Sequence 5, Application US/10054313
; Publication No. US20020110892A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0333
; CURRENT APPLICATION NUMBER: US/10/054,313
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-054-313-5

```

```

Query Match      62.0%; Score 959; DB 13; Length 216;
Best Local Similarity 70.2%; Pred. No. 5,7e-91;
Matches 177; Conservative 20; Mismatches 19; Indels 36; Gaps 1;

```

```

Qy 21 GSRGFGFYAVRRKRTGVFLTNWECRAQVDRPPAARFKKFPATEDEMAVFRKASPEVS 80
Db 1 GICGLGMFYAVRRGRRRPVEFLSWSECKAQVDRPPAARFKKFPATEDEMAVFRSSSPDGS 60
Qy 81 EGHENQGOSEAKPGKRLREPLDGDGHSQAQVPAKMKPSVEBAPPVSRDTFSYMGDFV 140
Db 61 KGOESAHQKSOAKTSRPREPL-----V 84
Qy 141 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAIEOAKTQ 200

```

```

Db 85 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAVMOAKAQ 144
Qy 201 NINKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 260
Db 145 NISKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 204
Qy 261 VPGHSGFTGNNE 272
Db 205 IPGHSGFVGNEE 216

```

```

RESULT 16
US-10-358-439-5
; Sequence 5, Application US/10358439
; Publication No. US20030144496A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H and Compositions And Uses Thereof
; FILE REFERENCE: ISPH0725
; CURRENT APPLICATION NUMBER: US/10/358,439
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/861,205
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/684,254
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/343,809
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 60/067,458
; PRIOR FILING DATE: 1997-12-04
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-358-439-5

```

```

Query Match      62.0%; Score 959; DB 14; Length 216;
Best Local Similarity 70.2%; Pred. No. 5,7e-91;
Matches 177; Conservative 20; Mismatches 19; Indels 36; Gaps 1;

```

```

Qy 21 GSRGFGFYAVRRKRTGVFLTNWECRAQVDRPPAARFKKFPATEDEMAVFRKASPEVS 80
Db 1 GICGLGMFYAVRRGRRRPVEFLSWSECKAQVDRPPAARFKKFPATEDEMAVFRSSSPDGS 60
Qy 81 EGHENQGOSEAKPGKRLREPLDGDGHSQAQVPAKMKPSVEBAPPVSRDTFSYMGDFV 140
Db 61 KGOESAHQKSOAKTSRPREPL-----V 84
Qy 141 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAIEOAKTQ 200
Db 85 VVYTDGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPGRQTNORAEITHAACKAVMOAKAQ 144
Qy 201 NINKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 260
Db 145 NISKLVLYTDSMFTINGITNWVQGMKKGWKTSAGKEVINKEDEVALERLTQGMIDIQMMH 204
Qy 261 VPGHSGFTGNNE 272
Db 205 IPGHSGFVGNEE 216

```

```

RESULT 17
US-10-358-439-2
; Sequence 2, Application US/10358439
; Publication No. US20030144496A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.

```

APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Human RNase H and Compositions And Uses Thereof  
FILE REFERENCE: ISPH0725  
CURRENT APPLICATION NUMBER: US/10/358,439  
PRIOR FILING DATE: 2003-02-03  
PRIOR APPLICATION NUMBER: 09/861,205  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: 09/684,254  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: 60/067,458  
PRIOR FILING DATE: 1997-12-04  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Gallus sp.  
US-10-358-439-2

Query Match 57.1%; Score 882; DB 14; Length 293;  
Best Local Similarity 57.7%; Pred. No. 8, 8e-83;  
Matches 173; Conservative 34; Mismatches 69; Indels 24; Gaps 5;

QY 1 MSWLLFLAHRVALALPCRRGSRGFMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 60  
DB 2 LKWL-----VALLSHSC-FVSKGGGMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 54  
QY 61 FATEDENAFVYKASPEVSEGENOHGOSEBAKPKRLREPLDGD-----GHSAQPY--- 114  
DB 55 FATEKEAMAFVQAGPPDQGSAPAEETHGASAVAQENASHREPEPTDVLCCNACKRREYEOS 114  
QY 115 -----AKMKPSVEBPAPVSRDTPSYMGDFVYVYTDGCCSSNGRRKPPAGIGVYWG 165  
DB 115 TNEEHTVRRAKH--DEOSTPVVSEAKFSYMEFVYVYTDGCCSSNGRRARAGIGVYWG 172  
QY 166 PGHPLNVGIRLPGRQTNQRAEITHAACKAIEQAKTONIKLVLYTDSMFTINGITWVQGM 225  
DB 173 PGHPLNISERLPGRQTNQRAEITHAACKAIEQAKSONIKLIIYTDSKFTINGITSVENW 232  
QY 226 KNGWMTSAGKEVINKEDFVALERLTQGMDIOWMHVPGHSGFTGNEADRLAREGASKOSE 285  
DB 233 KNGWMTSSGSGSVINKEDFQKLDLSKGIETIOWMHIPGHAGFGNEADRLAREGASKOK 292

RESULT 18  
US-09-861-205-2  
Sequence 2, Application US/09861205  
Patent No. US20020076712A1  
GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T.  
APPLICANT: Lima, Walter F.  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
FILE REFERENCE: ISPH-0333  
CURRENT APPLICATION NUMBER: US/09/861,205  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: 09/203,716  
PRIOR FILING DATE: 1998-12-02  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Gallus sp.  
US-09-861-205-2

Query Match 56.9%; Score 880; DB 9; Length 293;  
Best Local Similarity 57.3%; Pred. No. 1, 4e-82;  
Matches 172; Conservative 32; Mismatches 72; Indels 24; Gaps 4;

QY 1 MSWLLFLAHRVALALPCRRGSRGFMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 60  
DB 2 LKWL-----VALLSHSC-FVSKGGGMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 54  
QY 61 FATEDENAFVYKASPEVSEGENOHGOSEBAKPKRLREPLDGD-----GHSAQPY--- 106  
DB 55 FATEKEAMAFVQAGPPDQGSAPAEETHGASAVAQENASHREPEPTDVLCCNACKRREYEOS 114  
QY 107 -GHSAQPYAKMKPSVEBPAPVSRDTPSYMGDFVYVYTDGCCSSNGRRKPPAGIGVYWG 165  
DB 115 TNEEHTVRRAKH--DEOSTPVVSEAKFSYMEFVYVYTDGCCSSNGRRARAGIGVYWG 172  
QY 166 PGHPLNVGIRLPGRQTNQRAEITHAACKAIEQAKTONIKLVLYTDSMFTINGITWVQGM 225  
DB 173 PGHPLNISERLPGRQTNQRAEITHAACKAIEQAKSONIKLIIYTDSKFTINGITSVENW 232  
QY 226 KNGWMTSAGKEVINKEDFVALERLTQGMDIOWMHVPGHSGFTGNEADRLAREGASKOSE 285  
DB 233 KNGWMTSSGSGSVINKEDFQKLDLSKGIETIOWMHIPGHAGFGNEADRLAREGASKOK 292

RESULT 19  
US-10-054-313-2  
Sequence 2, Application US/10054313  
Publication No. US20020110892A1

GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T.  
APPLICANT: Lima, Walter F.  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
FILE REFERENCE: ISPH-0333  
CURRENT APPLICATION NUMBER: US/10/054,313  
PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/203,716  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-02  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Gallus sp.  
US-10-054-313-2

Query Match 56.9%; Score 880; DB 13; Length 293;  
Best Local Similarity 57.3%; Pred. No. 1, 4e-82;  
Matches 172; Conservative 32; Mismatches 72; Indels 24; Gaps 4;

QY 1 MSWLLFLAHRVALALPCRRGSRGFMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 60  
DB 2 LKWL-----VALLSHSC-FVSKGGGMFYAVRGRKTVFLTWNECRAQVDRFPAPRFXK 54  
QY 61 FATEDENAFVYKASPEVSEGENOHGOSEBAKPKRLREPLDGD-----GHSAQPY--- 106  
DB 55 FATEKEAMAFVQAGPPDQGSAPAEETHGASAVAQENASHREPEPTDVLCCNACKRREYEOS 114  
QY 107 -GHSAQPYAKMKPSVEBPAPVSRDTPSYMGDFVYVYTDGCCSSNGRRKPPAGIGVYWG 165  
DB 115 TNEEHTVRRAKH--DEOSTPVVSEAKFSYMEFVYVYTDGCCSSNGRRARAGIGVYWG 172  
QY 166 PGHPLNVGIRLPGRQTNQRAEITHAACKAIEQAKTONIKLVLYTDSMFTINGITWVQGM 225  
DB 173 PGHPLNISERLPGRQTNQRAEITHAACKAIEQAKSONIKLIIYTDSKFTINGITSVENW 232  
QY 226 KNGWMTSAGKEVINKEDFVALERLTQGMDIOWMHVPGHSGFTGNEADRLAREGASKOSE 285  
DB 233 KNGWMTSSGSGSVINKEDFQKLDLSKGIETIOWMHIPGHAGFGNEADRLAREGASKOK 292

RESULT 20  
US-10-262-511-84  
Sequence 84, Application US/10262511  
Publication No. US20040038223A1



```
GENERAL INFORMATION:
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: Peyman, John A.
APPLICANT: Kekuda, Rameesh
APPLICANT: Ju, Jingfang
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Paturajan, Meera
APPLICANT: Spytek, Kimberly A.
APPLICANT: Ellinger, Shlomit R.
APPLICANT: Ellerman, Karen
APPLICANT: Malysankar, Uriel M.
APPLICANT: Ort, Tatiana
APPLICANT: Gorman, Linda
APPLICANT: Zernusen, Bryan D.
APPLICANT: Anderson, David W.
APPLICANT: Zhong, Mei
APPLICANT: Catterton, Elina
APPLICANT: Ji, Weizhen
APPLICANT: Miller, Charles E.
APPLICANT: Raetelli, Luca
APPLICANT: Stone, David J.
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shimkels, Richard A.
APPLICANT: Rothenberg, Mark E.
APPLICANT: Leach, Martin D.
APPLICANT: Agee, Michele L.
APPLICANT: Berghs, Constance
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-462C
CURRENT APPLICATION NUMBER: US/10/262,511
CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/326,483
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/373,815
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/381,642
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 60/381,038
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/373,260
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/373,826
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,435
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: Curoseq1st version 0.1
SEQ ID NO 84
LENGTH: 152
TYPE: PRT
ORGANISM: Homo sapiens
US-10-262-511-84

Query Match          51.2%; Score 792; DB 12; Length 152;
Best Local Similarity 99.3%; Pred. No. 7.9e-74;
Matches 145; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 64 AKTONINKLVLYTDSMTFTINGITNWQMKNGKTSAGKEVINKEPVALERLTQGMDI 123
Qy 257 QMHWVPGHSGFIGNEDRLAREGAK 282
Db 124 QMHWVPGHSGFIGNEDRLAREGAK 149

RESULT 21
US-10-369-493-5383
Sequence 5383, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfang
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5383
LENGTH: 369
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5383

Query Match          28.5%; Score 440.5; DB 15; Length 369;
Best Local Similarity 30.0%; Pred. No. 8.9e-37;
Matches 108; Conservative 44; Mismatches 103; Indels 105; Gaps 6;

Qy 28 FYAVRGRKTVFLTWNECAOVDRPAPARFKKPEADEBAMFY--RK-----SASPEVS 80
Db 4 FYGVAHGFKRGVFTTEMAEAKQIDKFPQVYKKFETEEBAOKYVDKRPKGVSTFEEST 63
Qy 81 -----EGHENQHQ 89
Db 64 HDITYAVARGHSVGFNTNVEVKNHKNYRQPLHKKWSLTBEAIAVHYKYEGBEAKKA 123
Qy 90 ESEAKPGK--RLBPLDGDGHSQ-----PRAKMKPS 121
Db 124 ENEEKPDKSEKRSKRKADQESVEKKKKKFKFPAEAREKRVADNMSVPSKPTPA 183
Qy 122 VEPAPVSRDFTS-----YMGDFVVTYDGCSSNGRRKPRAGIGV 162
Db 184 VSTSSATRKRTHEGKTEKKMKTEBEVIDPEPANA PVYITDGCSSNGTKAKAGWV 243
Qy 163 YMGPHPLNVGIRLPGRTNQRAEIHAAKCAIEQAKTONINKLVLYTDSMTFTINGITNV 222
Db 244 YMGDDSEDNFEGPVYGAFTNNRGELIAVQKAIKALEKRLPKVYIKTDSULVLQSNMIWI 303
Qy 223 QGKNGKWKTSAGKEVINKEPVALERLTQGMDIQMHWVPGHSGFIGNEDRLAREGAK 282
Db 304 GHWKRGKWKTSSTGSEVLNQVLKMDLROKLVKFLHYVGHAGIDGENERADELARGAQ 363

RESULT 22
US-09-992-738-1
Sequence 1, Application US/0992738
Patent No. US20020160486A1
GENERAL INFORMATION:
APPLICANT: Wu, Hongjiang
APPLICANT: Lima, Walter F.
APPLICANT: Crooke, Stanley T.
TITLE OF INVENTION: Human RNase HI Mutants
FILE REFERENCE: ISPH-0614
CURRENT APPLICATION NUMBER: US/09/992,738
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/248,950
```



PRIOR FILING DATE: 2000-11-15  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 189  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-992-738-1

Query Match 27.5%; Score 424.5; DB 9; Length 189;  
 Best Local Similarity 53.5%; Pred. No. 1.6e-35;  
 Matches 151; Conservative 9; Mismatches 27; Indels 95; Gaps 37;

QY 1 MSMLFLAARVALAALPCRRGSGFPMFAVRRGRRTGYVLTWNECRAQVDRPPARFKK 60  
 DB 1 MSW---AARVAAA---CRRGSRG-GM-YAARRGRK-GV---WN-CRA-VDR--AAKKKA 43  
 QY 61 FATEDAMAFVRKSPASPEVSEGHENHOGSEAKPGKRLREPLDGDHESAOPYAKMKRP 120  
 DB 44 -----DAWA-VKRSAS---VSGH--NHG---SAKASKR-----RDGDGSHA---YAKMK- 81  
 QY 121 SUEPAPVRDPTFSYNGDVVVYTTDCCSSNGRRKPRAGICYYWGPHPLNGIRLPGRQ 180  
 DB 82 ----SVAVSRD--SYNGD-VVVY-DCCSSNGRRR-RAG-GYYWG-GH--NVG----- 121  
 QY 181 TNOBAEIIHAACKAIEAQTQNIKLVLTYDSMFTINGITNWQGMKKMGKTSAGKEVIN 240  
 DB 122 -RGRRAHAACRAAKKNTKY-----DSM--NG--NMV--GKKMGK-SAGK--VA 161  
 QY 241 KEDFVALERLTQMDIQMHVPHSGFIGNESADRLAREGAK 282  
 DB 162 KD-----VARGMD--WMHV-GHSG--GN--ADR--ARGAK 187

## RESULT 23

US-10-369-493-13271  
 ; Sequence 13271, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 13271  
 ; LENGTH: 300  
 ; TYPE: PRT  
 ; ORGANISM: Aspergillus nidulans  
 US-10-369-493-13271

Query Match 21.7%; Score 335; DB 15; Length 300;  
 Best Local Similarity 30.7%; Pred. No. 6.1e-26;  
 Matches 98; Conservative 37; Mismatches 98; Indels 86; Gaps 12;

QY 28 FYAVRRGRRTGYVLTWNECRAQVDRPPARFKKATEDAMAF--VRKSAS----- 76  
 DB 1 FYAIQGHGHPGYVTNANAOEQIRGFQKPKYKFSREBAEFCEARRGARSYCDPYRHQ 60  
 QY 77 -----EVSSEGHENHOG-----ESAKPGRLREPLD-GDGHESAOP 113  
 DB 61 TARGSLDERIYRPSRSGSVTRCGPFTTRCGGWASQCDP--RSDPARLSTRHOSRQ 118  
 QY 114 YAKMKPSVAPAPVSRDPTFSYNGDVVVYTTDCCSSNGRRKPRAGICYYWGPHPLNG 173  
 DB 119 LPHG---SOKDHPACS-----VETDSSLRNGRGVQAMAGVGVYFGPDGDSRFVS 164

QY 174 IRLP-----GRTNOBAEIIHAACKAIEAQTQNIKLVLTYDSMPT 214  
 DB 165 FRMSKYVLTITPSOQRLREPLKSGSRQTNORALTYILRALDIAFPH--RDYITVDSQYA 222  
 QY 215 INGITNWQGMKKMGKTSAGKEVINKEDEVALERLTQMD-----IQMHVPHG 264  
 DB 223 INCVTWFGKRSNNMVLTAADKKPVENKD---LVESILSKIDERTELRVKTLFEM--VKGH 277  
 QY 265 SGFTIGNEADRLAREGAKO 283  
 DB 278 DADPGNEADRLAVNGAQR 296

## RESULT 24

US-10-369-493-3170  
 ; Sequence 3170, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 3170  
 ; LENGTH: 275  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(275)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-369-493-3170

Query Match 20.1%; Score 310; DB 15; Length 275;  
 Best Local Similarity 30.1%; Pred. No. 2.1e-23;  
 Matches 88; Conservative 42; Mismatches 106; Indels 56; Gaps 8;

QY 28 FYAVRRGRRTGYVLTWNECRAQVDRPPARFKKATEDAMAFV--KSASPE--VS 80  
 DB 1 FYGVAVGRTPGYVTWMSIAQEVGVGMKPKYKFFTRAEBAEFVQWSGKPSPSRQGT 60  
 QY 81 EGHENHOGSEAKPGKRLREPLDGDHESAOPYAKMKPSVAPAPVSRDPTFSYNGDV 140  
 DB 61 SGRDXXPSHSCGSRASRKSRTKATGATOSAAQOAEALDDAKP-----V 105  
 QY 141 VVYTDGCCSSNGRRPRAGICYYWGPH-----PLANVIRL 176  
 DB 106 IYTDGSARGNGKGVAMAGVGVYFGGPREXVCSLTRKSKMCDSSXHTNRINSERL 165  
 QY 177 PCR-QTNOBAEIIHAACKAIEAQTQNIKLVLTYDSMFTINGITNWQGMKKMGKTS 234  
 DB 166 QGPVQTNGRAELTAVLRALAIPTQNCB---LRTDSQYITCVTSWYKMKMKNEWRNTK 222  
 QY 223 GGEVINKEPFLVLE-----RLTQMDIQMHVPHSGFIGNESADRLAREGA 281  
 DB 223 GEVSNODLIVAIRKKIDTRDKAETKFWVWGHGTDEGNIAADMLAVKA 274

## RESULT 25

US-10-369-493-2540  
 ; Sequence 2540, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 2540  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Schizosaccharomyces pombe  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(325)  
OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2540

Query Match 18.6%; Score 288; DB 15; length 325;

Best Local Similarity 24.4%; Pred. No. 5.2e-21;  
Matches 84; Conservative 42; Mismatches 104; Indels 114; Gaps 6;

QY 28 FYAVRGRKGTGVPLTWNECRA----- 48  
Db 8 YVAVARGRMTGIYETWMDCVSNGENGFVVKYKFHVKQVSIREYALLQFNLRFFQDDLLT 67  
QY 49 -----QVDRPAAAFKKFATEDENAFVR----- 72  
Db 68 FPKWHLVXQTLTNFYDEASDVYKGYGNNRYKKFDSYEAQEFCTEGSRYSSSSGPYR 127  
QY 73 -----KSASPEVSEGHENOHGE--SEAKPGKRLREPLDDGHESAQPYAKMKP 120  
Db 128 RSTTSYGYSPYSSSSSYSAHSKRYKKSISTEKDIEIFSNDRHKS----- 178  
QY 121 SVEBAPVSRDTFSYMGDFVVTYTGCCSSNGRRKPRAGIGVYWGPGHPLNVGIRLPG-R 179  
Db 179 -----IACSDRQVYVADGSSLRNGKKGAVAGCGVFPGNDPRNISVPLAGEE 225  
QY 180 QTNORAEIHAAKALIEQAKTONINKLVLYTDSMTFTINGITWVQGWKKNGKTSAGKEVI 239  
Db 226 QTNRAELQAILALENTS---GDLTIRSDSNYSIKSLTTLPLPKMKNDPKTSNQPVK 281  
QY 240 NKEDFVALERLTQGMDIOMMHVPGHSGFTGNEBADRLAREGAKQ 283  
Db 282 NLDLINRASDLMSPRNVSLVYKGHSTDYGNQADMLARRGASE 325

Search completed: October 7, 2004, 08:19:49  
Job time : 132 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:02:56 ; Search time 32 Seconds

(without alignments)  
461.407 Million cell updates/sec

Title: US-10-054-313-1

Perfect score: 1546

Sequence: 1 MSWFLHRLVLAALPCRR.....FIGNEADRLAREGAKQSD 286

Scoring table: BLOSUM62

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1546	100.0	286	3	US-09-203-716-1
2	1546	100.0	286	4	US-09-684-254-1
3	1546	100.0	286	4	US-09-409-926-1
4	959	62.0	216	3	US-09-203-716-5
5	959	62.0	216	4	US-09-684-254-5
6	959	62.0	216	4	US-09-409-926-4
7	882	57.1	293	4	US-09-409-926-2
8	880	56.9	293	3	US-09-203-716-2
9	880	56.9	293	4	US-09-684-254-2
10	272	17.6	441	4	US-09-540-236-3622
11	249.5	16.1	348	3	US-09-203-716-3
12	249.5	16.1	348	4	US-09-684-254-3
13	249.5	16.1	348	4	US-09-409-926-3
14	241	15.6	424	4	US-09-252-991A-30209
15	235.5	15.2	266	4	US-09-199-637A-157
16	217	14.0	502	4	US-09-328-352-5891
17	208.5	13.5	155	3	US-09-203-716-4
18	208.5	13.5	155	4	US-09-684-254-4
19	208.5	13.5	155	4	US-09-409-926-5
20	208.5	13.5	200	4	US-09-489-039A-12690
21	194.5	12.6	162	4	US-09-543-681A-6179
22	150	9.7	1079	2	US-08-929-967-8
23	150	9.7	1737	4	US-09-309-572-13
24	150	9.7	1737	4	US-09-718-096-13
25	148	9.6	665	2	US-08-929-967-7
26	144	9.3	1203	3	US-09-075-272-4
27	135.5	8.8	1031	4	US-08-811-682-15

28	132.5	8.6	3080	6	5223423-4	Patent No. 5223423
29	127.5	8.2	1055	2	US-08-659-251-5	Sequence 5, Appl
30	127.5	8.2	1055	3	US-09-256-490-5	Sequence 5, Appl
31	127.5	8.2	1055	5	PCT-US96-11445-5	Sequence 5, Appl
32	124.5	8.1	1016	4	US-09-206-551-46	Sequence 46, Appl
33	124.5	7.5	1016	4	US-09-625-972-24	Sequence 24, Appl
34	116.5	7.0	1014	4	US-09-319-588C-6	Sequence 6, Appl
35	107.5	6.9	1150	4	US-09-238-303-9	Sequence 9, Appl
36	107	6.9	1150	4	US-09-946-239-9	Sequence 9, Appl
37	105	6.8	1003	2	US-07-743-357-9	Sequence 9, Appl
38	105	6.8	1016	2	US-07-743-357-4	Sequence 4, Appl
39	102	6.6	1016	2	US-07-743-357-3	Sequence 3, Appl
40	102	6.6	1016	2	US-07-743-357-5	Sequence 5, Appl
41	101	6.5	1015	3	US-08-463-210-9	Sequence 9, Appl
42	101	6.5	1015	3	US-09-124-900-3	Sequence 3, Appl
43	101	6.5	1015	4	US-08-463-028-9	Sequence 9, Appl
44	101	6.5	1016	2	US-07-743-357-2	Sequence 2, Appl
45	100	6.5	560	4	US-09-752-652-1	Sequence 1, Appl
46	100	6.5	562	3	US-09-117-217-14	Sequence 14, Appl
47	100	6.5	562	4	US-09-735-487-14	Sequence 14, Appl
48	100	6.5	913	2	US-07-743-357-22	Sequence 22, Appl
49	100	6.5	1003	2	US-07-743-357-10	Sequence 10, Appl
50	100	6.5	1005	2	US-07-743-357-1	Sequence 1, Appl
51	98.5	6.4	654	4	US-08-979-847B-91	Sequence 91, Appl
52	98.5	6.4	683	4	US-08-979-847B-198	Sequence 198, Appl
53	98.5	6.4	683	4	US-08-979-847B-200	Sequence 200, Appl
54	98.5	6.4	683	4	US-08-979-847B-208	Sequence 208, Appl
55	98.5	6.4	683	4	US-08-979-847B-210	Sequence 210, Appl
56	98.5	6.4	768	4	US-08-979-847B-89	Sequence 89, Appl
57	98	6.3	913	2	US-07-743-357-6	Sequence 6, Appl
58	95	6.1	1003	2	US-07-743-357-7	Sequence 7, Appl
59	93	6.0	1004	2	US-07-743-357-8	Sequence 8, Appl
60	92	6.0	199	4	US-08-979-847B-92	Sequence 92, Appl
61	91	5.9	149	4	US-09-252-991A-23002	Sequence 23002, A
62	88.5	5.7	440	4	US-09-252-991A-26855	Sequence 26855, A
63	88.5	5.7	670	4	US-09-543-681A-7549	Sequence 7549, Ap
64	88	5.7	757	4	US-09-252-991A-23569	Sequence 23569, A
65	88	5.7	1003	4	US-09-309-572-17	Sequence 17, Appl
66	88	5.7	1003	4	US-09-718-096-17	Sequence 17, Appl
67	87.5	5.7	442	3	US-09-471-528-33	Sequence 33, Appl
68	87.5	5.7	442	3	US-09-634-530-33	Sequence 33, Appl
69	87.5	5.7	464	3	US-08-989-510A-8	Sequence 8, Appl
70	87.5	5.7	464	3	US-09-182-816-8	Sequence 8, Appl
71	87.5	5.7	464	3	US-09-182-816-11	Sequence 11, Appl
72	87.5	5.7	464	3	US-09-471-528-8	Sequence 8, Appl
73	87.5	5.7	464	3	US-09-471-528-11	Sequence 11, Appl
74	87.5	5.7	464	3	US-09-634-530-8	Sequence 8, Appl
75	87.5	5.7	464	3	US-09-634-530-11	Sequence 11, Appl
76	87.5	5.7	595	1	US-08-202-389-6	Sequence 6, Appl
77	87	5.6	822	4	US-09-029-755C-5	Sequence 5, Appl
78	85.5	5.5	402	4	US-09-252-991A-18436	Sequence 18436, A
79	84.5	5.5	288	4	US-09-252-991A-18436	Sequence 18436, A
80	84.5	5.5	631	1	US-08-202-389-8	Sequence 8, Appl
81	84.5	5.5	182	4	US-09-252-991A-21478	Sequence 21478, A
82	84	5.4	537	1	US-08-173-508-2	Sequence 2, Appl
83	84	5.4	537	2	US-08-265-310-2	Sequence 2, Appl
84	84	5.4	537	3	US-08-951-742-2	Sequence 2, Appl
85	84	5.4	735	4	US-09-252-991A-32172	Sequence 32172, A
86	83.5	5.4	735	4	US-09-252-991A-32172	Sequence 32172, A
87	83.5	5.4	414	4	US-09-252-991A-21676	Sequence 21676, A
88	83.5	5.4	573	4	US-09-252-991A-22802	Sequence 22802, A
89	83.5	5.4	843	4	US-09-417-197-117	Sequence 117, Appl
90	83.5	5.4	853	4	US-09-417-197-117	Sequence 117, Appl
91	83	5.4	369	4	US-09-252-991A-27583	Sequence 27583, A
92	83	5.4	386	4	US-09-252-991A-21635	Sequence 21635, A
93	83	5.4	804	4	US-09-252-991A-21676	Sequence 21676, A
94	82.5	5.3	298	4	US-09-252-991A-27658	Sequence 27658, A
95	82.5	5.3	849	4	US-09-252-991A-31525	Sequence 31525, A
96	82.5	5.3	1507	6	5268270-2	Patent No. 5268270
97	81.5	5.3	377	4	US-09-252-991A-31793	Sequence 31793, A
98	81.5	5.3	608	4	US-09-489-039A-13503	Sequence 13503, A
99	81	5.2	245	4	US-09-252-991A-32951	Sequence 32951, A
100	81	5.2	262	4	US-09-252-991A-19030	Sequence 19030, A

## ALIGNMENTS

```

RESULT 1
US-09-203-716-1
; Sequence 1, Application US/09203716
; Patent No. 6001653
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0333
; CURRENT APPLICATION NUMBER: US/09/203,716
; CURRENT FILING DATE: 1998-12-02
; EARLIER APPLICATION NUMBER: 60/067,458
; EARLIER FILING DATE: 1997-12-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-203-716-1

```

```

RESULT 2
US-09-684-254-1
; Sequence 1, Application US/09684254
; Patent No. 6376661
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0333
; CURRENT APPLICATION NUMBER: US/09/684,254
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-684-254-1

```

	Query Match	100.0%;	Score 1546;	DB 4;	Length 286;
	Best Local Similarity	100.0%;	Pred. No. 4, 2e-168;		
	Matches	286;	Conservative	0;	Mismatches 0;
				Indels	0;
				Gaps	0;
Qy	1	MSHLLFLAHVVALAALPCRGRSGRGFMFYAVRBGRKTVGLITNMCRAQVDRFPAPARKK	60		
Db	1	MSHLLFLAHRVALAALPCRGRSGRGFMFYARRRKRKTVGLITNMCRAQVDRFPAPARKK	60		
Qy	61	FATEDEAMAFVRKASPEVSESGHNOGQSEAKPKGLREPLDGDGHESAQPYAKHKMP	120		
Db	61	FATEDEAMAFVRKASPEVSESGHENQGQSEAKPKGLREPLDGDGHESAQPYAKHKMP	120		
Qy	121	SVEPAPVPSDPTFSYMDPVVVYVYTDGCCSSNGRKKPRAGITVWCGPHPLVNGIRLPERQ	180		
Db	121	SVEPAPVPSRTFSYMDPVVVYVYTDGCCSSNGRKKPRAGITVWCGPHPLVNGIRLPERQ	180		
Qy	181	TNCRAEITHAACKEAQAKTONINKLVLVYDSMPTINGITNNVQGMKKGMKTSACKEYIN	240		
Db	181	TNCRAEITHAACKEAQAKTONINKLVLVYDSMPTINGITNNVQGMKKGMKTSACKEYIN	240		
Qy	241	KEDVVALERLTQGMIDIOMHVPBGSGFGTGNENARLRAREGKQSED	286		
Db	241	KEDVVALERLTQGMIDIOMHVPBGSGFGTGNENARLRAREGKQSED	286		

Qy	1	MSWLLFLAHVVALAALPCRRGSRGFGMFYVARBRKQTGVFLTNHBCRAQVDRPPAARFKK	60
Db	1	MSWLLFLAHVVALAALPCRRGSRGFGMFYVARBRKQTGVFLTNHBCRAQVDRPPAARFKK	60
Qy	61	FATEDEAMAFVRKSASPEVSEGHENHQGSEAKPGRLREPLDGDGHSAQPYAKHMKP	120
Db	61	FATEDEAMAFVRKSASPEVSEGHENHQGSEAKPGRLREPLDGDGHSAQPYAKHMKP	120
Qy	121	SVBPAPVSRDTSYMDPFVYVYTDGCCSSNGRRKPPAGIGVMPGPHPLVNGIRLPGRQ	180
Db	121	SVBPAPVSRDTSYMDPFVYVYTDGCCSSNGRRKPPAGIGVMPGPHPLVNGIRLPGRQ	180
Qy	181	TNORAEIHAACKAIEAKTONINKLVLVTDSMPTINGITVMVQCKKONGMTSAGKEVIN	240
Db	181	TNORAEIHAACKAIEAKTONINKLVLVTDSMPTINGITVMVQCKKONGMTSAGKEVIN	240
Qy	241	KEBPVALERLTQCMDIQMHVPHGSGRTGNDEARLRAREGAKQSEED	286
Db	241	KEBPVALERLTQCMDIQMHVPHGSGRTGNDEARLRAREGAKQSEED	286



```

1  APPLICANT: Crooke, Stanley T.
2  APPLICANT: Lima, Walter F.
3  APPLICANT: Wu, Hongjiang
4  TITLE OF INVENTION: Human RNase H1 and Oligonucleotide Compositions Thereof
5  FILE REFERENCE: 1S1S4186
6  CURRENT APPLICATION NUMBER: US/09/409,926
7  CURRENT FILING DATE: 1999-09-30
8  NUMBER OF SEQ ID NOS: 33
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 2
11 LENGTH: 293
12 TYPE: PRT
13 ORGANISM: Artificial Sequence
14 FEATURE:
15 OTHER INFORMATION: Description of Artificial Sequence: No. 6617442el Sequence
16 US-09-409-926-2

```

Query Match	1%	Score	882	DB	4	Length	293
Best Local Similarity	57.7%	Pred. No.	3.2e-92				
Matches	173	Conservative	34	Mismatches	69	Indels	24
				Gaps	5		

  

QY	1	MSWLLFLAHVALAALPCRGRSGRGFMFYAVNRGKTVGLTWNECRPAQVDFEPAREKK	60
Db	2	LRMT-----VALLSHSC-FVSKGGGMFYAVNRGKGTGYRTMAACQOQVNNFPGASFEK	54
QY	61	FATEDEAWAFVRKSAPEVSEGEHQOESEAKRGKRLREPLDQ---GHSAPQY---	114
Db	55	FATEKEAFAFGAGPDDQGSAPARTHGASNAQENASRREPEFDVLCNACKRPYEOS	114
QY	115	-----AGMKPSEVPAPPVSDTTSYMGDFVYVYTTDCCSSNGRRKPPAGIGYTWG	165
Db	115	TNEEHTVRRARH--DEQSTPVSASAKSYMEGFVYVYTTDCCSSNGRRRARAAGIGYTWG	172
QY	166	PGHEPLNVGIRLPGRQTNRQRAEIIHAACKIAEQAKTQINIKLVLYTDSMFTINIGTMMVQGM	225
Db	173	PGHPLNISERLPGRQTNRQRAEIIHAACKIAEQAKSNIKLLIYTDSTKFTINIGTSMVEWM	232
QY	226	KKNQWKTSAAGEVINKEPVALERTQGMIDIQMHVPEHSGRTGNBEADRLAREGAKQSE	285
Db	233	KTNQGRITSSGSAVINKEPQKIDSTISKIGELQMHMIPGAGGQGNBEADRLAREGASQK	292

```

RESULT 8
US-09-203-716-2
: Sequence 2, Application US/09203716
: Patent No. 6001553
:
: GENERAL INFORMATION:
: APPLICANT: Crooke, Stanley T.
: APPLICANT: Lima, Walter F.
: APPLICANT: Wu, Hongjiang
: TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
: FILE REFERENCE: ISPH-0333
: CURRENT APPLICATION NUMBER: US/09/203,716
: CURRENT FILING DATE: 1998-12-02
: EARLIER APPLICATION NUMBER: 60/067,458
: EARLIER FILING DATE: 1997-12-04
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 2
: LENGTH: 293
:
: TYPE: PRT
:
: ORGANISM: Gallus sp.
:
: US-09-203-716-2

```

Query Match	56.9%	Score	880	DB 3:	Length	293
Best Local Similarity	57.3%	Pred. No.	5.5e-92			
Matches 172;	Conservative	32;	Mismatches	72;	Indels	24;
					Gaps	4

```

Qy      1  MSWLLFLAHRAVALALPCRGRSGRGGMVYAVRGRKGTGYFLTWNSCRAYQVREFPARFXK 60
        2  LRLW-----VALLSHSC-FVSKGGMGVYAVRKGKGTGYRTWAECCQQVNRFPASFXK 54
Qy      61  PATEDEAMAFVPRKASPEVSEGEHENQHQDESEAKGKRLLREFLDGD----- 106

```

```

Db      55  FATEENAAFPVAGCPDDQGSAPATHGASAAVAGENASHEREPEITDVLCCNACKRRYEOS 114
Qy      107 -GHESAOFPYAKMKESVEPAPPVSDTFSYMGDPVVVYTTDGCSSNGRRKPRAGIGVWG 165
Db      115  TNEEHTVARAKH--DEQSTPVVSEAKFSYMGEPFAVVYTTDGCSSNGRRARABAGIGVWG 172
Qy      166  PGHPLVNGVIRLPGRTNQRAETHAACKAIEQAKTONINKVLVYTDMSFTINGITVWVQGH 225
Db      173  PGHPLNIEERLEPGRTNQRAETHAACKAIEQAKSONIKLLIYVTSKFTINGITVWVWNN 232
Qy      226  KKGWKTASGKEVINKEDFVALERTLOGMDIQMHVPGHSGFTIGNEADRLAREGAKOSE 285
Db      223  KTNGRHTSSGGSVINKEDFQKLDSKIGIEIQMHVI PGHAGVQGNBEADRLAREGASKOK 292

RESULT 9
US-09-684-254-2
; Sequence 2, Application US/09684254
; Patent No. 6376661
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, HongJiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0333
; CURRENT APPLICATION NUMBER: US/09/684,254
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/203,716
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-684-254-2

```

```

Query Match      56.9%, Score 880; DB 4; Length 293;
Query Similarity 57.3%, Pred. No. 5.5e-92;
Best Local Match 172; Conservative 32; Mismatches 72; Indels 24; Gaps 4

QY      1 MSWILFLHRAVALAALPCRGRSGRGPFGMYAAYRGRKGTGFLTWNECRCAVDREPPARFPKK 60
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      2 LRLI-----VALLSHSC-FVSKGGGMYYAARKGQTGYRRYTWAECCQCVNRFPSAFYKK 54
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      61 FATEDEANAAYRKASPEVSEGHENQHQSSEAKGPKRLREPLDGD----- 106
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      55 FATEKEAMAFVAGGPPDQGSAPAEETHGASAVAGNASHREPERPDVLCNNAKRRYEQS 114
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      107 -GHSAQPYAKMKKSVEPAPVSDTTSYMGDFVYVYTTDCCSSNGRRKPPAGICVYWG 165
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      115 TNEETHVVRACH--DEQSTPVVSAKSKSYMEFPAVVYTDGCCSGNRRARAGITGVWG 172
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      166 PGHPLNVGILRPGROTNRAETHAACKIAEAKTKQINIKLVLYTDSMFTINGITWVQGM 225
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      173 PGHPLNISERLPGRQTNRAETHAACKIAEAKSQNIKKLIYTSKFTINIGITSVENVW 232
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      226 KKGNGKTSAGKEVINKEFVALERTLTGCGMDIOMVMVPGSGEFIGNEADRLRREGAKQSE 285
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      233 KTNNGRTISGGSVINKEDFQKIDLSISKGIETQMHI PGHAGVQGNBEADRLRREGASKKK 292
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 10  
US-09-540-236-3622  
; Sequence 3622, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS  
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2005-001  
 ; CURRENT APPLICATION NUMBER: US/09/540,236

CURRENT FILING DATE: 2000-04-04  
 NUMBER OF SEQ ID NOS: 3840  
 SEQ ID NO 3622  
 LENGTH: 441  
 TYPE: PRT  
 ORGANISM: M.cacarthalis  
 US-09-540-236-3622

Query Match 17.6%; Score 272; DB 4; Length 441;  
 Best Local Similarity 39.5%; Pred. No. 3e-22;  
 Matches 62; Conservative 20; Mismatches 51; Indels 24; Gaps 5;

136 MGDVYVYDGCSSNGRRKPRG-GIGY-----MGPGHPLANGIRLPGROTNQ 183  
 10 MSQALATDGCCKGKNGKGVLAHFNGBERLNG-GEF-----DTNN 59  
 184 RAHIAKCAIBQANTONINKLVLYTDSMTFTINGITNWQKKGKTSAGKEVINKED 243  
 60 RHELMALATLALATPAQ--IPQLMTDSGVKDGITOWIGWKLGKMKKADGKPVLNQDL 117  
 244 FVALERLTQGMIDQMHVPGHSGFIGNEDRLAREG 280  
 118 WORLDQTLTQNRITDQWIKHAGHAGNEMADQLANKG 154

RESULT 11  
 US-09-203-716-3  
 Sequence 3, Application US/09203716  
 Patent No. 6001653  
 GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T.  
 APPLICANT: Lima, Walter F.  
 APPLICANT: Wu, Hongjiang  
 TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
 FILE REFERENCE: ISPI-0333  
 CURRENT APPLICATION NUMBER: US/09/203, 716  
 CURRENT FILING DATE: 1998-12-02  
 EARLIER APPLICATION NUMBER: 60/067,458  
 EARLIER FILING DATE: 1997-12-04  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 3  
 LENGTH: 348  
 TYPE: PRT  
 ORGANISM: Saccharomyces sp.  
 US-09-203-716-3

Query Match 16.1%; Score 249.5; DB 3; Length 348;  
 Best Local Similarity 26.0%; Pred. No. 7.8e-20;  
 Matches 93; Conservative 38; Mismatches 112; Indels 115; Gaps 11;

26 GNFYAVRRGRKTGVFLTNNECPRAQVDRFPAPARKFKFATEDAMAFVRK-----S 74  
 5 GNFYAVRRGRKTGVFLTNNECPRAQVDRFPAPARKFKFATEDAMAFVRK-----S 74  
 75 ASPVSEGHENHOGSEAKPEKRLREPLDGDGHEA-----111  
 65 AGGQVSKPHHTQ-----KRVHRRNRLPHYSLSLTSSACSSLSANTNTFYSVKNV 117  
 112 -----QPYAKH-----MKPSVEPAP 127  
 118 IESKIFNNMKDQAVYKHGRGITFKFEDQLAENFISGMSADYKLMNISSEFSKTK 177  
 128 VSRDTFSYMGDFVVVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVG-IRLPGROTNQRAE 186  
 178 LSSNTM--YKSMNYYCDGSSFGNGTSSRAGYAGFEGAPBEINISEPLLSGQTNRAE 235  
 187 IHAACKAIBQ--AKTONINKLVLY--TDSMTFTINGITNWQKKGKTSAGKEV--- 239  
 236 IEAVSEALKKIWEKLTNEKEKYNVQIKTDSYVTKLNDRYMTYDNKKLEGLPNSDLIVP 295  
 240 -----NKEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKQ 283

DB 296 LVQRFVVKYKYYELNKECF---KNGKRFQIEW--VKGHDDPGNEMADFLAKGASR 347

RESULT 12  
 US-09-684-254-3  
 Sequence 3, Application US/09684254  
 Patent No. 6376661  
 GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T.  
 APPLICANT: Lima, Walter F.  
 APPLICANT: Wu, Hongjiang  
 TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
 FILE REFERENCE: ISPI-0333  
 CURRENT APPLICATION NUMBER: US/09/684,254  
 CURRENT FILING DATE: 09/203, 716  
 PRIOR FILING DATE: 1998-12-02  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 3  
 LENGTH: 348  
 TYPE: PRT  
 ORGANISM: Saccharomyces sp.  
 US-09-684-254-3

Query Match 16.1%; Score 249.5; DB 4; Length 348;  
 Best Local Similarity 26.0%; Pred. No. 7.8e-20;  
 Matches 93; Conservative 38; Mismatches 112; Indels 115; Gaps 11;

26 GNFYAVRRGRKTGVFLTNNECPRAQVDRFPAPARKFKFATEDAMAFVRK-----S 74  
 5 GNFYAVRRGRKTGVFLTNNECPRAQVDRFPAPARKFKFATEDAMAFVRK-----S 74  
 75 ASPVSEGHENHOGSEAKPEKRLREPLDGDGHEA-----111  
 65 AGGQVSKPHHTQ-----KRVHRRNRLPHYSLSLTSSACSSLSANTNTFYSVKNV 117  
 112 -----QPYAKH-----MKPSVEPAP 127  
 118 IESKIFNNMKDQAVYKHGRGITFKFEDQLAENFISGMSADYKLMNISSEFSKTK 177  
 128 VSRDTFSYMGDFVVVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVG-IRLPGROTNQRAE 186  
 178 LSSNTM--YKSMNYYCDGSSFGNGTSSRAGYAGFEGAPBEINISEPLLSGQTNRAE 235  
 187 IHAACKAIBQ--AKTONINKLVLY--TDSMTFTINGITNWQKKGKTSAGKEV--- 239  
 236 IEAVSEALKKIWEKLTNEKEKYNVQIKTDSYVTKLNDRYMTYDNKKLEGLPNSDLIVP 295  
 240 -----NKEDFVALERLTQGMIDQMHVPGHSGFIGNEDRLAREGAKQ 283

RESULT 13  
 US-09-409-926-3  
 Sequence 3, Application US/09409926  
 Patent No. 6617442  
 GENERAL INFORMATION:  
 APPLICANT: Crooke, Stanley T.  
 APPLICANT: Lima, Walter F.  
 APPLICANT: Wu, Hongjiang  
 TITLE OF INVENTION: Human RNase H and Oligonucleotide Compositions Thereof  
 FILE REFERENCE: ISIS4186  
 CURRENT APPLICATION NUMBER: US/09/409,926  
 CURRENT FILING DATE: 1999-09-30  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 348  
 TYPE: PRT  
 ORGANISM: Artificial sequence  
 FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: No. 6617442el Sequence  
US-09-409-926-3

Query Match 16.1%; Score 249.5; DB 4; Length 348;  
Best Local Similarity 26.0%; Pred. No. 7.8e-20;

Matches 93; Conservative 36; Mismatches 112; Indels 115; Gaps 11;

QY 26 GMEYAVRGRKGTGVLFWNECRQVDFRPAARFKKATDEDAVFAVRK-----S 74

DB 5 GNFYAVKRGRTGIVTNECKQVDDGAYKKFNSYQAKSFLGQPTTSYSGSSTH 64

QY 75 ASPEVSGHENGQSEAPKRLREPLDGDGHEA----- 111

DB 65 AGQVSKPHTTQ-----KRVHRNRPVLYSLTSSACSSLSANTNTFYSVKSNPN 117

QY 112 -----OPYAKH-----MKPSVEPAPP 127

DB 118 IBSKIFNNMDCQAYVHKKGITPKKEDQLAENFISGMSADHYKLMNISKSFSKTK 177

QY 128 VSRDTSYMGDFVVTYDGCSSNGRRKPPAGIGVYVPGHPVNG-IRLPGRQTNRAR 186

DB 178 LSNNTM--YKSNMNVYCDGSSFGNGTSSRAGYAFEGAPBEERISEPLLSGAQTNRAR 235

QY 187 IHAACKAIEQ-AKTQINLVLY--TDSMFTINGITTMVQKKKGMKTSAGKEYI-- 239

DB 236 IEAVSEALKKIWEKLTNEKERVNQITDSEYVTKLNDRYMTYDNKKLEGLPNSDLIVP 295

QY 240 -----NKEDFVALERLTQGMIDQMHVPGHSGFIGNEADRLARBAKQ 283

DB 296 LVQRFVVKRYEVLNECF-----KNGKFOLEW--VGHGDGDPGENADFLAKKASR 347

RESULT 14

US-09-252-991A-30209

Sequence 30209, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252.991A

PRIOR FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 30209

LENGTH: 424

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30209

Query Match 15.6%; Score 241; DB 4; Length 424;  
Best Local Similarity 25.8%; Pred. No. 9.9e-19;

Matches 92; Conservative 36; Mismatches 126; Indels 102; Gaps 14;

QY 9 HRVALAALPCRGSRGFMFYAVRGRKGTGVLFWNECRQVDFRPAAR---FKKATFD 65

DB 87 HRADDSAR--CAPGAAAGRGHRLRGR-----LAAQRTRCG--RGPAAKTPGFLPVASPS 137

QY 66 EAVA-FVRKASPEVSEGHENH-----GOSEAKPKRLREPLDGDGHE 109

DB 138 PARSRAVRSGRPPADRHQPMPEVGHPSLFRGCLAPGLHSSVAGRLAOPAGLAGE 197

QY 110 SA-----QPYAKHMK-----PSVEAPPVSRPT-- 132

DB 198 TALRVLSAACVGSALAPSGFALGRRLAVFGRRLLSIGTQAGRCVAVABEQARTAR 257

QY 133 -----FSYMG-----DFVVVYTDGCCSSNGRRKPPAGIGV-----WGP 167

DB 258 SAGAHGSGESQPARFRVLGNTDKQVVIYTDGACKGNPGRGAGLALTYGAREELMG-G 316

QY 168 HPLNVGIRLPGRTQNRARIEAACKAIEQAKTQINLVLYTDSMFTINGITTMVQGMK 227

DB 317 EF-----DTNNRMELMAIQLALAKRSCPRLI--TDSVVMGITEMLEPNWK 365

QY 228 NGWKTSGKEVINKPEFVALERLTQGMIDQMHVPGHSGFIGNEADRLARBAKQ 283

DB 366 RGWKTASKQPVKNADLMQALDEQVABHGVEMQVVRGHTDPPGNERADLANRGVAB 421

RESULT 15

US-09-199-637A-157

Sequence 157, Application US/09199637A

Patent No. 6355411

GENERAL INFORMATION:

APPLICANT: Ausubel, Frederick

APPLICANT: Goodman, Howard M.

APPLICANT: Rahme, Laurence G.

APPLICANT: Mahajan-Miklos, Shailina

APPLICANT: Cao, Hui

APPLICANT: Drenkard, Eliana

APPLICANT: Teongalle, John

TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID

FILE REFERENCE: 00786/361002

CURRENT APPLICATION NUMBER: US/09/199,637A

CURRENT FILING DATE: 1998-11-25

PRIOR FILING DATE: 1997-11-25

NUMBER OF SEQ ID NOS: 437

SOFTWARE: FaastSeq for Windows Version 4.0

SEQ ID NO 157

LENGTH: 266

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-157

Query Match 15.2%; Score 235.5; DB 4; Length 266;  
Best Local Similarity 31.6%; Pred. No. 2.1e-18;

Matches 61; Conservative 29; Mismatches 74; Indels 29; Gaps 6;

QY 102 PLDGDGHEAQPAYAKMKRSVEPAPVSRDTSYMG---DFVVVYTDGCCSSNGRRKPP 157

DB 89 PVAPQEARARSAGAGGSGSPA-----RFRNLGMDTKQVVIYTDACKGNPGRGSM 142

QY 158 AGIGVY-----WGHGHPVNVGIRLPGRTQNRARIEAACKAIEQAKTQINLVLYTD 210

DB 143 GALLTYKGAERELMG-GEF-----DTNNRMELMAIQLALAKRSCPRLI--TD 190

QY 211 SMFTINGITTMVQGMKNGKMTSAGKEVINKPEFVALERLTQGMIDQMHVPGHSGFIGN 270

DB 191 SEVVMGITEMLEPNWKRGKKTASKQPVKNADLMQALDQVABHGVEMQVVRGHTDPPGN 250

QY 271 EADRLARBAKQ 283

DB 251 ERADOLANRGVAB 263

RESULT 16

US-09-328-352-5891

Sequence 5891, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 5891

LENGTH: 502



TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-5891

Query Match 14.0%; Score 217; DB 4; Length 502;  
Best Local Similarity 31.5%; Pred. No. 7e-16;  
Matches 52; Conservative 24; Mismatches 53; Indels 36; Gaps 6;

QY 133 PSYMDPVVVTDDGCCSSNGRRKPRAGIGVYWGPR-----GHPLVNGIRLPKR 179  
DB 41 FRFMQITTLVYDGCGRN-----PGLG-GWGAIVITEQGHKLFGGPR-----D 84  
QY 180 QNNOAABIHAACKAEQAKTQINIKLVYTDMSFTINGITNNVQKKGWTSAGKEVI 239  
DB 85 TTNNMELTALAEISFPCPD--AQLIVTDSNYKQGITTEIRHKKKNM-----KDYK 137  
QY 240 NKEDEVALERLTQGMIDQMMHVPKSGFISGNEADRLAREGAKOS 284  
DB 138 NPDLMQKLDVACAGRNIEKNWIKGHAGHGAGNEMADQLANIGADKT 182

RESULT 17  
US-09-203-716-4  
Sequence 4, Application US/09203716  
Patent No. 6001653

GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T.  
APPLICANT: Lima, Walter F.  
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
FILE REFERENCE: ISPH-0333  
CURRENT APPLICATION NUMBER: US/09/203, 716  
CURRENT FILING DATE: 1998-12-02  
EARLIER APPLICATION NUMBER: 60/067,458  
EARLIER FILING DATE: 1997-12-04  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 155  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-203-716-4

Query Match 13.5%; Score 208.5; DB 3; Length 155;  
Best Local Similarity 36.6%; Pred. No. 1.1e-15;  
Matches 52; Conservative 16; Mismatches 69; Indels 5; Gaps 2;

QY 140 VVVYTDGCCSSNGRRKPRAGIGVYWGPRHPLNVGIRLPGRQTNORAEIHAACKAIEQAKT 199  
DB 5 VEIITDSCLGMPGPGYCALIRYGRKRTFSAGYT---RTNNMELMAAIVALEALKE 61  
QY 200 QNINKLVYTDMSFTINGITNNVQKKGWTSAGKEVINKEDEVALERLTQGMIDQMM 259  
DB 62 H--CEVILSTDSQYVRQGITQWIIHMKKKGWTKADKPKVKNVDLMQRLDALGQHQIKWE 119  
QY 260 HYPGHSFGFISGNEADRLAREGA 281  
DB 120 WKGHAGHPNERCDELARAAA 141

RESULT 18  
US-09-684-254-4  
Sequence 4, Application US/09684254  
Patent No. 6376661

GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T.  
APPLICANT: Lima, Walter F.  
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof  
FILE REFERENCE: ISPH-0333  
CURRENT APPLICATION NUMBER: US/09/684, 254  
CURRENT FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 09/203, 716

PRIOR FILING DATE: 1998-12-02  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 155  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-684-254-4

Query Match 13.5%; Score 208.5; DB 4; Length 155;  
Best Local Similarity 36.6%; Pred. No. 1.1e-15;  
Matches 52; Conservative 16; Mismatches 69; Indels 5; Gaps 2;

QY 140 VVVYTDGCCSSNGRRKPRAGIGVYWGPRHPLNVGIRLPGRQTNORAEIHAACKAIEQAKT 199  
DB 5 VEIITDSCLGMPGPGYCALIRYGRKRTFSAGYT---RTNNMELMAAIVALEALKE 61  
QY 200 QNINKLVYTDMSFTINGITNNVQKKGWTSAGKEVINKEDEVALERLTQGMIDQMM 259  
DB 62 H--CEVILSTDSQYVRQGITQWIIHMKKKGWTKADKPKVKNVDLMQRLDALGQHQIKWE 119  
QY 260 HYPGHSFGFISGNEADRLAREGA 281  
DB 120 WKGHAGHPNERCDELARAAA 141

RESULT 19  
US-09-409-926-5  
Sequence 5, Application US/09409926  
Patent No. 6617442

GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T.  
APPLICANT: Lima, Walter F.  
TITLE OF INVENTION: Human RNase H1 and Oligonucleotide Compositions Thereof  
FILE REFERENCE: ISIS4186  
CURRENT APPLICATION NUMBER: US/09/409, 926  
CURRENT FILING DATE: 1999-09-30  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 155  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE: Description of Artificial Sequence: No. 6617442el Sequence  
US-09-409-926-5

Query Match 13.5%; Score 208.5; DB 4; Length 155;  
Best Local Similarity 36.6%; Pred. No. 1.1e-15;  
Matches 52; Conservative 16; Mismatches 69; Indels 5; Gaps 2;

QY 140 VVVYTDGCCSSNGRRKPRAGIGVYWGPRHPLNVGIRLPGRQTNORAEIHAACKAIEQAKT 199  
DB 5 VEIITDSCLGMPGPGYCALIRYGRKRTFSAGYT---RTNNMELMAAIVALEALKE 61  
QY 200 QNINKLVYTDMSFTINGITNNVQKKGWTSAGKEVINKEDEVALERLTQGMIDQMM 259  
DB 62 H--CEVILSTDSQYVRQGITQWIIHMKKKGWTKADKPKVKNVDLMQRLDALGQHQIKWE 119  
QY 260 HYPGHSFGFISGNEADRLAREGA 281  
DB 120 WKGHAGHPNERCDELARAAA 141

RESULT 20  
US-09-489-039A-12690  
Sequence 12690, Application US/09489039A  
Patent No. 6610836

GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS



Query Match 9.7%; Score 150; DB 4; Length 1737;  
 Best Local Similarity 28.5%; Pred. No. 2.1e-07;  
 Matches 61; Conservative 33; Mismatches 78; Indels 42; Gaps 12;

QY OHGQSEAKPGKRLREPLDGDHESAOPIYAK--HMKPSVEPAP-PVSRDTSYMGDFVVV 142  
 DB 1129 QGPPVALNPAITLPLPREGLOHNCIDILAEHAGTRPDLTQPLPDADHTW----- 1179  
 QY 143 YTDG-CCSSNGRRKPRAGI---GYWGPGLNVLGIRLPGKQTNORAEIHAACKAIEQA 197  
 DB 1180 YTDGSSLLQEGQRKAGAAVTETEVIMAKA-----LPAGTSAQRAELIALTOALKMA 1231  
 QY 198 KTONINKLVLYTDSMFTINGITNWVG--WKNKGKTSAGKEVINKEDEPVALER---LTQ 252  
 DB 1232 EGKRLN---VYDTSRYAF--ATAHIGETIYRRRGLITSEGEIKKKDETLALKALFLPK 1286  
 QY 253 GMDIQMHVPGH-----SGFIGNBEADRLAREGA 281  
 DB 1287 RLST--IHCPGHQKHSABARGNMADQAAKRAA 1318

RESULT 24  
 US-09-718-096-13  
 ; Sequence 13, Application US/09718096  
 ; Patent No. 6589763

GENERAL INFORMATION:  
 APPLICANT: Von Laer, Melke-Dorothee  
 TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV  
 FILE REFERENCE: 35-195  
 CURRENT APPLICATION NUMBER: US/09/718, 096  
 PRIOR FILING DATE: 1998-11-22  
 PRIOR APPLICATION NUMBER: DE 19856463.5  
 PRIOR FILING DATE: 1998-11-26  
 PRIOR APPLICATION NUMBER: EP 99250415.9  
 PRIOR FILING DATE: 1999-11-25  
 PRIOR APPLICATION NUMBER: US 09/309,572  
 PRIOR FILING DATE: 1999-05-11  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 13  
 LENGTH: 1737  
 TYPE: PRT  
 ORGANISM: Moloney murine leukemia virus  
 FEATURE:  
 OTHER INFORMATION: gag-pol protein  
 US-09-718-096-13

Query Match 9.7%; Score 150; DB 4; Length 1737;  
 Best Local Similarity 28.5%; Pred. No. 2.1e-07;  
 Matches 61; Conservative 33; Mismatches 78; Indels 42; Gaps 12;

QY 86 OHGQSEAKPGKRLREPLDGDHESAOPIYAK--HMKPSVEPAP-PVSRDTSYMGDFVVV 142  
 DB 1129 QGPPVALNPAITLPLPREGLOHNCIDILAEHAGTRPDLTQPLPDADHTW----- 1179  
 QY 143 YTDG-CCSSNGRRKPRAGI---GYWGPGLNVLGIRLPGKQTNORAEIHAACKAIEQA 197  
 DB 1180 YTDGSSLLQEGQRKAGAAVTETEVIMAKA-----LPAGTSAQRAELIALTOALKMA 1231  
 QY 198 KTONINKLVLYTDSMFTINGITNWVG--WKNKGKTSAGKEVINKEDEPVALER---LTQ 252  
 DB 1232 EGKRLN---VYDTSRYAF--ATAHIGETIYRRRGLITSEGEIKKKDETLALKALFLPK 1286  
 QY 253 GMDIQMHVPGH-----SGFIGNBEADRLAREGA 281  
 DB 1287 RLST--IHCPGHQKHSABARGNMADQAAKRAA 1318

RESULT 25  
 US-08-929-967-7  
 ; Sequence 7, Application US/08929967  
 ; Patent No. 5891637

GENERAL INFORMATION:

APPLICANT: Ruppert, Siegfried J.W.  
 TITLE OF INVENTION: Construction of Full-length cDNA Libraries  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatIn (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/929,967  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Dreger, Ginger R.  
 REGISTRATION NUMBER: 33,055  
 REFERENCE/DOCKET NUMBER: P1035R1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-3216  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 665 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-929-967-7

Query Match 9.6%; Score 148; DB 2; Length 665;  
 Best Local Similarity 27.4%; Pred. No. 8.2e-08;  
 Matches 59; Conservative 32; Mismatches 80; Indels 44; Gaps 11;

QY 86 OHGQSEAKPGKRLREPLDGDHES---ADPYAGHMKPSVEPAPVSRDTSYMGDFVV 141  
 DB 472 QGPPVALNPAITLPLPREGLOHNCIDILAEHAGTRSDLTQPLPDADH-----T 521  
 QY 142 YTDGCC-SSNGRRKPRAGI---GYWGPGLNVLGIRLPGKQTNORAEIHAACKAIEQA 196  
 DB 522 YTTDSSFLQEGQRKAGAAVTETEVIMAKA-----LPAGTSAQRAELIALTOALKMA 573  
 QY 197 AKTONINKLVLYTDSMFTINGITNWVG--WKNKGKTSAGKEVINKEDEPVALER---LT 251  
 DB 574 EGKRLN---VYDTSRYAF--ATAHIGETIYRRRGLITSEGEIKKKDETLALKALFLPK 628  
 QY 253 GMDIQMHVPGH-----SGFIGNBEADRLAREGA 281  
 DB 629 RLST--IHCPGHQKHSABARGNMADQAAKRAA 661

Search completed: October 7, 2004, 08:08:23  
 Job time : 35 secs

This page blank (sep10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 08:02:54 : Search time 127 Seconds

(without alignment) 636.288 Million cell updates/sec

Title: US2000054-313-1

Perfect score: 1546

Sequence: 1 MSWFLFLHVRALALAPCRF.....FTGNEDRLAREGAKQSD 286

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : A.Geneseq23Jan04:\*

1: geneseqp19808:\*\*\*  
2: geneseqp19808:\*\*\*  
3: geneseqp20009:\*\*\*  
4: geneseqp20018:\*\*\*  
5: geneseqp20026:\*\*\*  
6: geneseqp2003as:\*\*\*  
7: geneseqp2003bs:\*\*\*  
8: geneseqp2004s:\*\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1546	100.0	286	2	AA25094 Human typ
2	1546	100.0	286	4	AB97508 Human typ
3	1546	100.0	286	5	ABP53036 Human typ
4	1526	98.7	286	5	ABP53037 Human typ
5	1523	98.5	286	3	AA770235 Human RNA
6	1522	98.4	286	5	ABP53038 Human RNA
7	1522	98.4	286	6	ADA05726 Human NOV
8	1518	98.2	286	5	ABP53039 Human RNA
9	1513	97.9	286	5	ABP53039 Human RNA
10	1510	97.7	286	5	ABP53039 Human RNA
11	1508	97.5	286	5	ABP53039 Human RNA
12	1508	97.5	286	5	ABP53039 Human RNA
13	1501	97.1	286	5	ABP53039 Human RNA
14	1489	96.3	286	5	ABP53039 Human RNA
15	1489	96.3	286	5	ABP53039 Human RNA
16	1489	96.3	286	5	ABP53039 Human RNA
17	1489	96.3	286	5	ABP53039 Human RNA
18	1489	96.3	286	5	ABP53039 Human RNA
19	1489	96.3	286	5	ABP53039 Human RNA
20	1489	96.3	286	5	ABP53039 Human RNA
21	1489	96.3	286	5	ABP53039 Human RNA
22	1489	96.3	286	5	ABP53039 Human RNA
23	1489	96.3	286	5	ABP53039 Human RNA
24	1489	96.3	286	5	ABP53039 Human RNA
25	1489	96.3	286	5	ABP53039 Human RNA

26	249.5	16.1	348	2	AA25096 Yeast typ
27	249.5	16.1	348	6	AB97510 Yeast typ
28	249.5	16.1	348	6	ABP53681 Protein s
29	235.5	15.3	154	4	ABP53681 Protein s
30	235.5	15.3	154	4	ABP53681 Protein s
31	234.5	15.2	145	5	ABP79988 N. gonorr
32	217	14.0	502	6	ADA34604 Actinobac
33	213.5	13.7	155	2	AA779152 E.coli ri
34	211.5	13.6	155	2	AA779152 E.coli ri
35	209.5	13.5	155	2	AA779152 E.coli ri
36	208.5	13.5	155	2	AA779152 E.coli ri
37	208.5	13.5	155	2	AA779152 E.coli ri
38	205.5	13.3	140	6	AA027111 LINE rebr
39	205.5	13.3	155	2	AA779152 E.coli ri
40	204.5	13.2	155	2	AA779152 E.coli ri
41	200.5	13.0	157	6	ABP53039 Human RNA
42	190.5	12.3	197	6	ABP53039 Human RNA
43	190.5	12.3	208	6	ABP53039 Human RNA
44	173.5	11.2	164	2	AA779152 E.coli ri
45	172.5	11.2	309	6	AA779152 E.coli ri
46	172.5	11.2	309	6	AA779152 E.coli ri
47	155	10.0	1784	2	AA779152 E.coli ri
48	155	10.0	1784	2	AA779152 E.coli ri
49	152	9.8	1194	2	AA779152 E.coli ri
50	150	9.7	1079	2	AA779152 E.coli ri
51	150	9.7	1199	3	AA779152 E.coli ri
52	150	9.7	1224	2	AA779152 E.coli ri
53	150	9.7	1737	3	AA779152 E.coli ri
54	149	9.6	1145	4	AA779152 E.coli ri
55	149	9.6	1145	4	AA779152 E.coli ri
56	148.5	9.6	1035	2	AA779152 E.coli ri
57	145	9.4	672	3	AA779152 E.coli ri
58	144	9.3	1197	4	AA779152 E.coli ri
59	144	9.3	1203	2	AA779152 E.coli ri
60	140.5	9.1	1124	2	AA779152 E.coli ri
61	140	9.1	1196	2	AA779152 E.coli ri
62	137	8.9	683	5	AA779152 E.coli ri
63	137	8.9	683	5	AA779152 E.coli ri
64	137	8.9	683	5	AA779152 E.coli ri
65	137	8.9	683	5	AA779152 E.coli ri
66	137	8.9	683	5	AA779152 E.coli ri
67	137	8.9	683	5	AA779152 E.coli ri
68	137	8.9	683	5	AA779152 E.coli ri
69	137	8.9	683	5	AA779152 E.coli ri
70	137	8.9	683	5	AA779152 E.coli ri
71	137	8.9	683	5	AA779152 E.coli ri
72	137	8.9	683	5	AA779152 E.coli ri
73	137	8.9	683	5	AA779152 E.coli ri
74	137	8.9	683	5	AA779152 E.coli ri
75	137	8.9	683	5	AA779152 E.coli ri
76	137	8.9	683	5	AA779152 E.coli ri
77	137	8.9	683	5	AA779152 E.coli ri
78	137	8.9	683	5	AA779152 E.coli ri
79	137	8.9	683	5	AA779152 E.coli ri
80	137	8.9	683	5	AA779152 E.coli ri
81	135.5	8.8	1031	5	AA779152 E.coli ri
82	135.5	8.8	1124	2	AA779152 E.coli ri
83	134	8.7	1193	2	AA779152 E.coli ri
84	132.5	8.6	3080	1	AA779152 E.coli ri
85	132.5	8.6	3210	1	AA779152 E.coli ri
86	130.5	8.4	1015	4	AA779152 E.coli ri
87	129.5	8.4	1035	7	AA779152 E.coli ri
88	128.5	8.3	1053	2	AA779152 E.coli ri
89	128	8.3	1105	1	AA779152 E.coli ri
90	127.5	8.2	1055	2	AA779152 E.coli ri
91	127.5	8.2	1055	2	AA779152 E.coli ri
92	127.5	8.2	1055	2	AA779152 E.coli ri
93	125.5	8.1	1064	1	AA779152 E.coli ri
94	125.5	8.1	1027	1	AA779152 E.coli ri
95	125.5	8.1	1036	1	AA779152 E.coli ri
96	125.5	8.1	1036	1	AA779152 E.coli ri
97	125.5	8.1	1056	1	AA779152 E.coli ri
98	125.5	8.1	1056	1	AA779152 E.coli ri

99 124.5 8.1 1018 3 AAB01769  
100 124 8.0 742 4 ABG19846

Aab01769 Simian im  
ABG19846 Novel hum

## ALIGNMENTS

RESULT 1  
ID AAY25094 standard; protein; 286 AA.

XX AAY25094;

DT 24-AUG-1999 (first entry)

DE Human type 2 RNase H protein.

XX Type 2 RNase H; oligonucleotide-RNA duplex; cleavage; antisense therapy;  
KM interaction; target mRNA; human.

XX Homo sapiens.

XX WO928447-A1.

PD 10-JUN-1999.

XX 02-DEC-1998; 98WO-US025488.

XX 04-DEC-1997; 97US-0067458P.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 1999-394857/33.

PT New isolated human RNase H polypeptide and nucleic acid, used to develop  
PT agents for enhancing the efficiency of antisense therapy.

PS Claim 1; Page 28-29; 37pp; English.

XX This invention describes a novel human RNase H polypeptide and nucleic  
CC acid which can be used to cleave the RNA strand of oligonucleotide-RNA  
CC duplexes and to develop agents for enhancing the efficiency of antisense  
CC therapy. The products of the invention are useful in defining the  
CC interaction of human Type 2 RNase H and antisense oligonucleotides and  
CC identifying methods for enhancing this interaction so that antisense  
CC oligonucleotides are more effective at inhibiting their target mRNA. The  
CC products can be used for enhancing the efficacy of antisense  
CC oligonucleotide therapies. This sequence represents the human type 2  
CC RNase H used in the method of the invention

XX Sequence 286 AA;

Query Match 100.0%; Score 1546; DB 2; Length 286;

Best Local Similarity 100.0%; Pred. No. 1.5e-148;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFGMFYAVRGRKTGVFLTNNECRAOYDRPPAARFKK 60  
DB 1 MSWLLFLAHRVALAALPCRRGSRGFGMFYAVRGRKTGVFLTNNECRAOYDRPPAARFKK 60  
QY 61 FATEDAMAFVRKASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAOPYAKMKP 120  
DB 61 FATEDAMAFVRKASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAOPYAKMKP 120  
QY 121 SVBPAPVSRDTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVGIRLPGRQ 180  
DB 121 SVBPAPVSRDTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVGIRLPGRQ 180  
QY 181 TNORAEIHAACKAIEOAKTONINKLVLYTDSMTFINGITNMVQGMKNGMKTSGAKEVIN 240  
DB 181 TNORAEIHAACKAIEOAKTONINKLVLYTDSMTFINGITNMVQGMKNGMKTSGAKEVIN 240

QY 241 KEDFVALERLTQGMDIQMHVPGHSGFIGNEBADRILAREGAKOSED 286  
DB 241 KEDFVALERLTQGMDIQMHVPGHSGFIGNEBADRILAREGAKOSED 286

RESULT 2  
ID AAB97508 standard; protein; 286 AA.

XX AAB97508;

DT 14-AUG-2001 (first entry)

DE Human type II RNase H protein.

XX Human; RNase H type II; RNase H1 cleavage substrate; antisense therapy;  
KM gene therapy.

XX Homo sapiens.

XX WO200123613-A1.

PD 05-APR-2001.

XX 23-SEP-2000; 2000WO-US026729.

XX 30-SEP-1999; 99US-00409926.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H, Manoharan M;

XX WPI; 2001-343164/36.

PT Chimeric oligonucleotides that can serve as substrates for human RNase  
PT H1, useful for enhancing the effectiveness of antisense gene therapies.

PS Disclosure; Fig 6; 178pp; English.

XX The present invention provides a number of DNA-RNA oligonucleotides which  
CC can act as substrates for human RNase H1 (a type II RNase). The sequence  
CC consists of two portions, one of which is capable of supporting cleavage  
CC of a complementary target RNA and the other of which is incapable of  
CC supporting such cleavage. These can be used to enhance the effectiveness  
CC of antisense therapies. The present sequence is human RNase H1

XX Sequence 286 AA;

Query Match 100.0%; Score 1546; DB 4; Length 286;

Best Local Similarity 100.0%; Pred. No. 1.5e-148;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFGMFYAVRGRKTGVFLTNNECRAOYDRPPAARFKK 60  
DB 1 MSWLLFLAHRVALAALPCRRGSRGFGMFYAVRGRKTGVFLTNNECRAOYDRPPAARFKK 60  
QY 61 FATEDAMAFVRKASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAOPYAKMKP 120  
DB 61 FATEDAMAFVRKASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAOPYAKMKP 120  
QY 121 SVBPAPVSRDTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVGIRLPGRQ 180  
DB 121 SVBPAPVSRDTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYWGPGHPLNVGIRLPGRQ 180  
QY 181 TNORAEIHAACKAIEOAKTONINKLVLYTDSMTFINGITNMVQGMKNGMKTSGAKEVIN 240  
DB 181 TNORAEIHAACKAIEOAKTONINKLVLYTDSMTFINGITNMVQGMKNGMKTSGAKEVIN 240  
QY 241 KEDFVALERLTQGMDIQMHVPGHSGFIGNEBADRILAREGAKOSED 286  
DB 241 KEDFVALERLTQGMDIQMHVPGHSGFIGNEBADRILAREGAKOSED 286

## RESULT 3

ABP53035  
ID ABP53036 standard; protein, 286 AA.XX  
AC ABP53036;

DT 06-NOV-2002 (first entry)

XX Human type 2 RNase H protein sequence SEQ ID NO:6.

XX RNase H; antisense technology; inhibition.

XX Homo sapiens.

XX WO200264841-A1.

XX 22-AUG-2002.

XX 12-FEB-2002; 2002WO-US004243.

XX 12-FEB-2001; 2001US-00781712.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 2002-657606/70.

Use of a mammalian, particularly human, RNase H, for treating an animal with a disease or condition associated with a human RNase H, for inhibiting the expression of a protein, or for reducing cellular RNA via antisense technology.

Claim 8; Page 53-54; 70pp; English.

The present invention describes a method for promoting the inhibition of the expression of a protein comprising employing a mammalian RNase H polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA complex duplex occurs. Also described is a compound 8 to 50 nucleobases in length targeted to the nucleic acid encoding the human RNase HII polypeptide, where the compound specifically hybridizes with and inhibits the expression of a human RNase HII polypeptide. The compound, which is an antisense oligonucleotide, is useful for inhibiting the expression of a human RNase HII polypeptide in cells or tissues, as well as for treating an animal with a disease or condition associated with a human RNase HII polypeptide. The method is useful for inhibiting the expression of a protein, particularly for reducing cellular RNA via antisense technology. The present sequence represents human type 2 RNase H given in the exemplification of the present invention

XX Sequence 286 AA;

Query Match 100.0%; Score 1546; DB 5; Length 286;

Best Local Similarity 100.0%; Pred. No. 1.5e-148; Mismatches 0; Indels 0; Gaps 0;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALALPCRRSGRGFMFYAVRRGRKTGVFLTWNECRAQVDFPAPARFXK 60  
DB 1 MSWLLFLAHRVALALPCRRSGRGFMFYAVRRGRKTGVFLTWNECRAQVDFPAPARFXK 60  
QY 61 FATEDEANAFVYKSKSPVSEBHENOHGSESAKPKRLREPLDGDGHSAQPYAKHMKP 120  
DB 61 FATEDEANAFVYKSKSPVSEBHENOHGSESAKPKRLREPLDGDGHSAQPYAKHMKP 120  
QY 121 SVEPAPVPSRDTFSYMGDFVYVYTTGCCSSNGRRPRAGIGYWGPGHPLANGIRLPGRQ 180  
DB 121 SVEPAPVPSRDTFSYMGDFVYVYTTGCCSSNGRRPRAGIGYWGPGHPLANGIRLPGRQ 180  
QY 181 TNGRAEIHAAKCAIEQAKTQINIKLVLYTDSMFTTNGITNMTVOGKCKGKMTSAGKEVIN 240  
DB 181 TNGRAEIHAAKCAIEQAKTQINIKLVLYTDSMFTTNGITNMTVOGKCKGKMTSAGKEVIN 240

QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFIGNEADRLAREGAKQSED 286  
DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFIGNEADRLAREGAKQSED 286

RESULT 4  
ABP53037  
ID ABP53037 standard; protein, 286 AA.XX  
AC ABP53037;

DT 06-NOV-2002 (first entry)

XX Human RNase H protein sequence SEQ ID NO:7.

XX RNase H; antisense technology; inhibition.

XX Homo sapiens.

XX WO200264841-A1.

XX 22-AUG-2002.

XX 12-FEB-2002; 2002WO-US004243.

XX 12-FEB-2001; 2001US-00781712.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 2002-657606/70.

Use of a mammalian, particularly human, RNase H, for treating an animal with a disease or condition associated with a human RNase H, for inhibiting the expression of a protein, or for reducing cellular RNA via antisense technology.

Claim 8; Page 55; 70pp; English.

The present invention describes a method for promoting the inhibition of the expression of a protein comprising employing a mammalian RNase H polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA complex duplex occurs. Also described is a compound 8 to 50 nucleobases in length targeted to the nucleic acid encoding the human RNase HII polypeptide, where the compound specifically hybridizes with and inhibits the expression of a human RNase HII polypeptide. The compound, which is an antisense oligonucleotide, is useful for inhibiting the expression of a human RNase HII polypeptide in cells or tissues, as well as for treating an animal with a disease or condition associated with a human RNase HII polypeptide. The method is useful for inhibiting the expression of a protein, particularly for reducing cellular RNA via antisense technology. The present sequence represents a human RNase H protein sequence, given in the exemplification of the present invention

XX Sequence 286 AA;

Query Match 98.7%; Score 1526; DB 5; Length 286;

Best Local Similarity 98.6%; Pred. No. 1.7e-146; Mismatches 2; Indels 0; Gaps 0;

Matches 282; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSWLLFLAHRVALALPCRRSGRGFMFYAVRRGRKTGVFLTWNECRAQVDFPAPARFXK 60  
DB 1 MSWLLFLAHRVALALPCRRSGRGFMFYAVRRGRKTGVFLTWNECRAQVDFPAPARFXK 60  
QY 61 FATEDEANAFVYKSKSPVSEBHENOHGSESAKPKRLREPLDGDGHSAQPYAKHMKP 120  
DB 61 FATEDEANAFVYKSKSPVSEBHENOHGSESAKPKRLREPLDGDGHSAQPYAKHMKP 120  
QY 121 SVEPAPVPSRDTFSYMGDFVYVYTTGCCSSNGRRPRAGIGYWGPGHPLANGIRLPGRQ 180  
DB 121 SVEPAPVPSRDTFSYMGDFVYVYTTGCCSSNGRRPRAGIGYWGPGHPLANGIRLPGRQ 180  
QY 181 TNGRAEIHAAKCAIEQAKTQINIKLVLYTDSMFTTNGITNMTVOGKCKGKMTSAGKEVIN 240  
DB 181 TNGRAEIHAAKCAIEQAKTQINIKLVLYTDSMFTTNGITNMTVOGKCKGKMTSAGKEVIN 240

QY	181	TNGBRAETHAACRKAIEQAKTQNIKLVLVYDSMFTNGITNNVQGWKXGWTSSAGKEVIN	240
Db	181	TNGBRAETHAACRKAIEQAKTQNIKLVLVYDSMFTNGITNNVQGWKXGWTSSAGKEVIN <td>240</td>	240
QY	241	KEPVALERLTQGMDIQMHVPGHSGFTONEADRLAREGAKQSED	286
Db	241	KEPVALERLTQGMDIQMHVPGHSGFTONEADRLAREGAKQSED	286
XX	AAV70235	standard; protein; 286 AA.	
XX	AAV70235		
XX	AAV70235		
XX	06-JUN-2000	(first entry)	
DE	Human RNA-associated protein-16 (RNAAP-16).		
XX	RNA-associated protein; RNAAP; human; clone 2073417; cytosolic;		
KW	immunosuppressive; antiinflammatory; keratolytic; neuroprotective;		
KW	antitumor; antileukemic; hepatotropic; antiparasitic; virucide; anti-HIV;		
KW	antiallergic; antirheumatic; antiarthritic; ophthalmological; autoimmune;		
KW	actinocidal; cell proliferative disorder; inflammation; cirrhosis;		
KW	actinic keratosis; burns; arteriosclerosis; artherosclerosis;		
KW	hepatitis; myelofibrosis; primary thrombocythemia; psoriasis; cancer;		
KW	mixed connective tissue disease; MCTD; HIV; uveitis; Crohn's disease;		
KW	allergy; rheumatoid arthritis; parasitic infection.		
XX	Homo sapiens.		
OS			
XX			
XX	Key	Location/Qualifiers	
XX	Peptide	1..18	
XX	FT	/label= Signal_peptide	
XX	FT	19..286	
XX	Protein	/label= Mature_human_RNA_associated_protein-16	
XX	Modified-site	22	
FT	Modified-site	/note= "Potential phosphorylation site"	
FT	Modified-site	42	
FT	Modified-site	/note= "Potential phosphorylation site"	
FT	Modified-site	63	
FT	Modified-site	/note= "Potential phosphorylation site"	
FT	Modified-site	96	
FT	Modified-site	/note= "Potential phosphorylation site"	
FT	Region	137..282	
XX		/note= "Signature sequence of Rnae H"	
XX	WO200011171-A2.		
XX	02-MAR-2000.		
XX	20-AUG-1999;	99WO-US019361.	
XX	21-AUG-1998;	98US-0097550P.	
XX	12-JAN-1999;	99US-0115639P.	
XX	(INCYTE PHARM INC.		
XX	Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA,		
XX	Patterson C, Baughn MR, Lal P, Bandman O, Reddy R, Azimzal Y,		
XX	Shih L, Yang Y, Lu DM,		
XX	WPI; 2000-237651/20.		
XX	N-PSDB; AAZ51265.		
XX	Human RNA-associated proteins useful in diagnosing, treating and		
XX	preventing cell proliferative, autoimmune, inflammatory and infectious		
XX	disorders.		
XX	Claim 1; Page 95-96; 123pp; English.		
XX	The present amino acid sequence is the human RNA-associated protein-16		
XX	(RNAAP-16), identified in Incyte clone 2073417, derived from ISLTN0701		

Query Match	98.5%	Score 1523;	DB 3;	Length 286;
Best Local Similarity	98.3%	Fred. No. 3.3e-146;		
Matches 281;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1 MSWLLFLHRAVALALPCRGRSGRGFMFVAVRRGRKTVFLTNCECAQVDRFPAPARPK	60		
DB	1 MSWLLFLHRAVALALPCRGRSGRGFMFVAVRRGRKTVFLTNCECAQVDRFPAPARPK	60		
QY	61 FATEDMAAFVFKASPEVSGHENOGESBAKPKRLRPLDGDGHSAQPIAKMKP	120		
DB	61 FATEDMAAFVFKASPEVSGHENOGESBAKPKRLRPLDGDGHSAQPIAKMKP	120		
QY	121 SVEPAPVSRPTFSYMGDFVYVYTDGCCSSNGRRKPRAGIGVYMGPHPLVNGIRLPQRQ	180		
DB	121 SVEPAPVSRPTFSYMGDFVYVYTDGCCSSNGRRKPRAGIGVYMGPHPLVNGIRLPQRQ	180		
QY	181 TNGRAEIHAAKAEIQAKTQNIKLVTYDTSMTFTINGITNNVQGKXGKMTSAGKEVIN	240		
DB	181 TNGRAEIHAAKAEIQAKTQNIKLVTYDTSMTFTINGITNNVQGKXGKMTSAGKEVIN	240		
QY	241 KEDVALERLTQGMIDQMMHYPGHSGLFGNEADRLAREGAKQSED	286		
DB	241 KEDVALERLTQGMIDQMMHYPGHSGLFGNEADRLAREGAKQSED	286		
RESULT 6				
ABP53038	ABP53038 standard; protein: 286 AA.			
XX	AC	ABP53038;		
XX	DT	06-NOV-2002 (first entry)		
XX	DE	Human RNase HI protein sequence SEQ ID NO:8.		
XX	KM	RNase H; antisense technology; inhibition.		
XX	OS	Homo sapiens.		
XX	PN	WO200264841-A1.		
XX	PD	22-AUG-2002.		
XX	PF	12-FEB-2002; 2002WO-US004243.		
XX	PR	12-FEB-2001; 2001US-00781712.		
XX	PA	(ISIS-) ISIS PHARM INC.		
XX	PI	Crooke ST, Lima WF, Wu H;		
XX	DR	WPI; 2002-657606/70.		
PT	Use of a mammalian, particularly human, RNase H, for treating an animal with a disease or condition associated with a human RNase H, for inhibiting the expression of a protein, or for reducing cellular RNA via antisense technology.			



XX Claim 8; Page 56-57; 70pp; English.  
 PS The present invention describes a method for promoting the inhibition of  
 CC the expression of a protein comprising employing a mammalian RNase H  
 CC polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA  
 CC complex duplex occurs. Also described is a compound 8 to 50 nucleobases  
 CC in length targeted to the nucleic acid encoding the human RNase H1  
 CC polypeptide, where the compound specifically hybridizes with and inhibits  
 CC the expression of a human RNase H1 polypeptide. The compound, which is  
 CC an antisense oligonucleotide, is useful for inhibiting the expression of  
 CC a human RNase H1 polypeptide in cells or tissues, as well as for  
 CC treating a animal with a disease or condition associated with a human  
 CC RNase H1 polypeptide. The method is useful for inhibiting the expression  
 CC of a protein, particularly for reducing cellular RNA via antisense  
 CC technology. The present sequence represents human RNase H1 given in the  
 CC exemplification of the present invention

SQ Sequence 286 AA;

Query Match 98.4%; Score 1532; DB 5; Length 286;  
 Best Local Similarity 98.3%; Pred. No. 4, 2e-146;  
 Matches 281; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSWFLFLARVALALPCRGRSGFMFYAVRGRGTGVTLMNCRQAQVDFPAPARFK 60  
 Db 1 MSWFLFLARVALALPCRGRSGFMFYAVRGRGTGVTLMNCRQAQVDFPAPARFK 60  
 QY 61 FATEDENAFVFKASPEVSEGHENQGESEAKPKRLREPLDGDGHEAQPVAKMKP 120  
 Db 61 FATEDENAFVFKASPEVSEGHENQGESEAKPKRLREPLDGDGHEAQPVAKMKP 120  
 QY 121 SVEPAPSPDPTFSMDPQVYVYTGCCSSNGRRPRAGIGYVGRGHLNGLPGRQ 180  
 Db 121 SVEPAPSPDPTFSMDPQVYVYTGCCSSNGRRPRAGIGYVGRGHLNGLPGRQ 180  
 QY 181 TNORAEIHAAKAIQAKTONINKLVLYTDSMTFTNGITNVOGWKMGKTSAGKEVIN 240  
 Db 181 TNORAEIHAAKAIQAKTONINKLVLYTDSMTFTNGITNVOGWKMGKTSAGKEVIN 240  
 QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRIAREGAKQSED 286  
 Db 241 KEDFVALERLTQGMIDIQMHVPGHSGFTIGNEADRIAREGAKQSED 286

RESULT 7  
 ADA05726  
 ID ADA05726 standard; protein; 286 AA.

XX ADA05726;  
 AC ADA05726;  
 XX 06-NOV-2003 (first entry)  
 DT Human NOVA5e protein SEQ ID NO:86.  
 DE  
 XX human; NOVA; antidiabetic; anorectic; antibacterial; vinuclide;  
 KW immunomodulator; cytoskeletal; neurotropic; neuroprotective;  
 KW antiparkinsonian; antilipemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 OS  
 XX WO2003029424-A2.  
 PN  
 XX 10-APR-2003.  
 PD  
 XX 02-OCT-2002; 2002WO-US031373.  
 PF  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0331058P.  
 PR 24-OCT-2001; 2001US-033266P.  
 PR 24-OCT-2001; 2001US-0345629P.  
 PR 29-OCT-2001; 2001US-0345757P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373848P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Sathuraj G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Paturnajon M, Seytek KA, Edinger SR, Ellerman K, Malvankar UM;  
 PI Ort T, Gotman L, Zernhusen BD, Anderson DW, Zhong M, Catereron B;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CER, Shenoy SG;  
 PI Shinkes RA, Rothenberg ME, Leach MD, Agee ML, Berghe C, Dipippo VA;  
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
 XX MPI: 2003-381626/36.  
 DR N-PSDB; ADA05725.  
 XX  
 PT New NOVA polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVA-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 PS  
 XX Claim 1; Page 161; 586pp; English.  
 PS  
 XX The present invention describes NOVA proteins, where X can be 1 to 55  
 CC (e.g. NOVA). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVA protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVA  
 CC sequences have antidiabetic, anorectic, antibacterial, vinuclide,  
 CC immunomodulator, cytoskeletal, neurotropic, neuroprotective, antiparkinsonian  
 CC and antilipemic activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX Sequence 286 AA;

Query Match 98.4%; Score 1522; DB 6; Length 286;  
 Best Local Similarity 98.3%; Pred. No. 4.2e-146;  
 Matches 281; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSWLFLLAHRVALAALPCRRSGRGFGMFYAVRRGRKTGVFLTNWECRAQVDRPPARFKK 60  
 DB 1 MSWFLFLAHRVALAALPCRRSGRGFGMFYAVRRGRKTGVFLTNWECRAQVDRPPARFKK 60  
 QY 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPKGRLEPLDGDGHESAOPYAKMKP 120  
 DB 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPKGRLEPLDGDGHESAOPEYAKMKP 120  
 QY 121 SVBPAPVSRDFTSYMGDFVYVYTTDCCSSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
 DB 121 SVBPAPVSRDFTSYMGDFVYVYTTDCCSSNGRRRPRAGIGYVWGPHPLNVGIRLPGRQ 180  
 QY 181 TNGRAEIIHAAKCAIEQAKTONINKLVLYTDSMTINGITNWVQGMKNGKTSAGKEVIN 240  
 DB 181 TNGRAEIIHAAKCAIEQAKTONINKLVLYTDSMTINGITNWVQGMKNGKTSAGKEVIN 240  
 QY 241 KEPFVALERTQGMIDQMMHVPGHSGFTIGNEADRLAREGAKOSED 286  
 DB 241 KEPFVALERTQGMIDQMMHVPGHSGFTIGNEADRLAREGAKOSED 286

#### RESULT 8

ABP53039  
 ID ABP53039 standard; protein; 286 AA.

XX ABP53039;

DT 06-NOV-2002 (first entry)

XX Human RNase HII protein sequence SEQ ID NO:9.

XX Human RNase H; antisense technology; inhibition.

XX Homo sapiens.

XX WO200264841-A1.

XX 22-AUG-2002.

XX 12-FEB-2002; 2002WO-US004243.

XX 12-FEB-2001; 2001US-00781712.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 2002-657606/70.

XX Use of a mammalian, particularly human, RNase H, for treating an animal  
 PT with a disease or condition associated with a human RNase H, for  
 PT inhibiting the expression of a protein, or for reducing cellular RNA via  
 PT antisense technology.

XX Claim 8; Page 57-58; 70pp; English.

XX The present invention describes a method for promoting the inhibition of  
 CC the expression of a protein comprising employing a mammalian RNase H

CC polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA  
 CC complex duplex occurs. Also described is a compound 8 to 50 nucleobases  
 CC in length targeted to the nucleic acid encoding the human RNase HII  
 CC polypeptide, where the compound specifically hybridises with and inhibits  
 CC the expression of a human RNase HII polypeptide. The compound, which is  
 CC an antisense oligonucleotide, is useful for inhibiting the expression of  
 CC a human RNase HII polypeptide in cells or tissues, as well as for  
 CC treating an animal with a disease or condition associated with a human  
 CC RNase HII polypeptide. The method is useful for inhibiting the expression  
 CC of a protein, particularly for reducing cellular RNA via antisense  
 CC technology. The present sequence represents a human RNase HII protein  
 CC sequence, given in the exemplification of the present invention

XX Sequence 286 AA;

Query Match 98.2%; Score 1518; DB 5; Length 286;  
 Best Local Similarity 97.9%; Pred. No. 1.1e-145;  
 Matches 280; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSWLFLLAHRVALAALPCRRSGRGFGMFYAVRRGRKTGVFLTNWECRAQVDRPPARFKK 60  
 DB 1 MSWFLFLAHRVALAALPCRRSGRGFGMFYAVRRGRKTGVFLTNWECRAQVDRPPARFKK 60  
 QY 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPKGRLEPLDGDGHESAOPYAKMKP 120  
 DB 61 FATEDAMAFVRKSASPEVSEGHENHGOSEAKPKGRLEPLDGDGHESAOPEYAKMKP 120  
 QY 121 SVBPAPVSRDFTSYMGDFVYVYTTDCCSSNGRRKPRAGIGYVWGPHPLNVGIRLPGRQ 180  
 DB 121 SVBPAPVSRDFTSYMGDFVYVYTTDCCSSNGRRRPRAGIGYVWGPHPLNVGIRLPGRQ 180  
 QY 181 TNGRAEIIHAAKCAIEQAKTONINKLVLYTDSMTINGITNWVQGMKNGKTSAGKEVIN 240  
 DB 181 TNGRAEIIHAAKCAIEQAKTONINKLVLYTDSMTINGITNWVQGMKNGKTSAGKEVIN 240  
 QY 241 KEPFVALERTQGMIDQMMHVPGHSGFTIGNEADRLAREGAKOSED 286  
 DB 241 KEPFVALERTQGMIDQMMHVPGHSGFTIGNEADRLAREGAKOSED 286

#### RESULT 9

ABB83371  
 ID ABB83371 standard; protein; 286 AA.

XX ABB83371;

DT 04-SEP-2002 (first entry)

XX Human wild-type RNase H1.

XX Human; RNase H1; enzyme.

XX Homo sapiens.

XX Location/Qualifiers

FT 1..73 /label= Region\_I

FT /note= "Contains the double-stranded RNA (dsRNA) -binding motif"

FT 74..135 /label= Region\_II

FT 136..286 /label= Region\_III

FT /note= "Included within this region are the conserved  
 FT amino acid residues that form the catalytic site, the  
 FT divalent cation binding site, and the basic substrate-  
 FT binding domain"

XX WO200240635-A2.

XX 23-MAY-2002.

XX 14-NOV-2001; 2001WO-US043929.

XX 15-NOV-2000; 2000US-0248950P.  
 PR (ISIS-) ISIS PHARM INC.  
 PA  
 XX Mu H, Lima WF, Crooke ST;  
 XX WPI; 2002-519372/55.  
 DR  
 XX Novel human RNase H1 polypeptide comprising mutations compared to wild  
 PT type human RNase H1 useful for inhibiting expression of selected protein  
 PT by antisense oligonucleotide targeted to RNA encoding selected protein.  
 XX  
 XX Disclosure; Fig 1; 39pp; English.  
 PS  
 CC The present sequence is the wild-type human RNase H1 sequence. RNase H  
 CC hydrolyses RNA in RNA-DNA hybrids, and the activity of RNase H1 is Mg+2  
 CC dependent and inhibited by both Mn+2 and the sulphhydryl blocking agent N-  
 CC ethylmaleimide. The amino acid sequence of human RNase H1 displays strong  
 CC homology with RNase H1 from yeast, chicken, E. coli and mouse. This  
 CC sequence was used to generate RNase H1 mutants (see ABB83372-ABB83376) by  
 CC site-directed mutagenesis. By studying the mutants generated in the  
 CC present invention, it was found that although the conserved amino acid  
 CC residues of the putative catalytic site and basic substrate-binding  
 CC domain are required for RNase H activity, deletion of either the  
 CC catalytic site or the basic substrate-binding domain did not ablate  
 CC binding to the heteroduplex substrate  
 CC  
 XX Sequence 286 AA;  
 SQ  
 Query Match 97.9%; Score 1513; DB 5; Length 286;  
 Best Local Similarity 98.3%; Pred. No. 3 5e-145;  
 Matches 281; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MSWLLFLAHRVALALPCRGRSGFGMFYAVRGRKTVFLTNWCRAQVDRFPAPARFK 60  
 DB 1 MSWLLFLAHRVALALPCRGRSGFGMFYAVRGRKTVFLTNWCRAQVDRFPAPARFK 60  
 QY 61 FATEDENAFVFKKSPVSEBGENOHQSESAKSKRLREPLDDGHSAPYAKHMKP 120  
 DB 61 FATEDENAFVFKKSPVSEBGENOHQSESAKSKRLREPLDDGHSAPYAKHMKP 120  
 QY 121 SVEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
 QY 181 TNORAEIHAAKCAIQAKTONINKLVLYTDSMTFTNGITNNVQGGKGMKTSAGKEVIN 240  
 DB 181 TNORAEIHAAKCAIQAKTONINKLVLYTDSMTFTNGITNNVQGGKGMKTSAGKEVIN 240  
 QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTGNBEADRLARAGAKOSD 286  
 DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFTGNBEADRLARAGAKOSD 286  
 RESULT 10  
 ABB83374  
 ID ABB83374 standard; protein; 286 AA.  
 XX  
 XX ABB83374;  
 XX AC  
 XX 04-SEP-2002 (first entry)  
 XX DT  
 XX Mutant RNase H1, E186Q.  
 XX DE  
 XX Human; RNase H1; enzyme; mutant.  
 XX KM  
 XX Homo sapiens.  
 XX OS  
 XX Synthetic.  
 XX FH  
 XX Key Location/Qualifiers  
 FT Region 1..73  
 FT /label= Region\_1

FT /note= "Contains the double-stranded RNA (dsRNA) -binding  
 FT motif"  
 FT 74..135  
 FT Region  
 FT /label= Region\_II  
 FT 136..286  
 FT Region  
 FT /label= Region\_III  
 FT /note= "Included within this region are the conserved  
 FT amino acid residues that form the catalytic site, the  
 FT divalent cation binding site, and the basic substrate-  
 FT binding domain"  
 FT 186  
 FT /note= "Glu substituted by Gln"  
 XX  
 XX WO200240635-A2.  
 XX PD  
 XX 23-MAY-2002.  
 XX PF  
 XX 14-NOV-2001; 2001WO-US043929.  
 XX PR  
 XX 15-NOV-2000; 2000US-0248950P.  
 XX PA  
 XX (ISIS-) ISIS PHARM INC.  
 XX PI  
 XX Mu H, Lima WF, Crooke ST;  
 XX WPI; 2002-519372/55.  
 DR  
 XX Novel human RNase H1 polypeptide comprising mutations compared to wild  
 PT type human RNase H1 useful for inhibiting expression of selected protein  
 PT by antisense oligonucleotide targeted to RNA encoding selected protein.  
 XX  
 XX Disclosure; Page; 39pp; English.  
 PS  
 CC The present sequence is a mutant RNase H1 sequence. RNase H hydrolyses  
 CC RNA in RNA-DNA hybrids, and the activity of RNase H1 is Mg+2 dependent  
 CC and inhibited by both Mn+2 and the sulphhydryl blocking agent N-  
 CC ethylmaleimide. The amino acid sequence of human RNase H1 displays strong  
 CC homology with RNase H1 from yeast, chicken, E. coli and mouse. This  
 CC sequence was generated from the wild-type human RNase H1 sequence  
 CC (ABB83371) by site-directed mutagenesis. It was found that although the  
 CC conserved amino acid residues of the putative catalytic site and basic  
 CC substrate-binding domain are required for RNase H activity, deletion of  
 CC either the catalytic site or the basic substrate-binding domain did not  
 CC ablate binding to the heteroduplex substrate. This mutant exhibited an  
 CC undetectable rate of cleavage of a heteroduplex substrate, i.e. the  
 CC mutation ablated the cleavage activity of the enzyme. Note: The present  
 CC sequence was not shown in the specification, but was derived from the  
 CC wild-type human RNase H1 sequence shown on Fig 1  
 CC  
 XX Sequence 286 AA;  
 SQ  
 Query Match 97.7%; Score 1510; DB 5; Length 286;  
 Best Local Similarity 97.9%; Pred. No. 7e-145;  
 Matches 280; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MSWLLFLAHRVALALPCRGRSGFGMFYAVRGRKTVFLTNWCRAQVDRFPAPARFK 60  
 DB 1 MSWLLFLAHRVALALPCRGRSGFGMFYAVRGRKTVFLTNWCRAQVDRFPAPARFK 60  
 QY 61 FATEDENAFVFKKSPVSEBGENOHQSESAKSKRLREPLDDGHSAPYAKHMKP 120  
 DB 61 FATEDENAFVFKKSPVSEBGENOHQSESAKSKRLREPLDDGHSAPYAKHMKP 120  
 QY 121 SVEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRQ 180  
 QY 181 TNORAEIHAAKCAIQAKTONINKLVLYTDSMTFTNGITNNVQGGKGMKTSAGKEVIN 240  
 DB 181 TNORAEIHAAKCAIQAKTONINKLVLYTDSMTFTNGITNNVQGGKGMKTSAGKEVIN 240  
 QY 241 KEDFVALERLTQGMIDIQMHVPGHSGFTGNBEADRLARAGAKOSD 286  
 DB 241 KEDFVALERLTQGMIDIQMHVPGHSGFTGNBEADRLARAGAKOSD 286



PS Disclosure; Page: 39pp; English.

XX The present sequence is a mutant RNase H1 sequence. RNase H hydrolyses  
CC RNA in RNA-DNA hybrids, and the activity of RNase H1 is Mg<sup>2+</sup> dependent  
CC and inhibited by both Mn<sup>2+</sup> and the sulphhydryl blocking agent N-  
CC ethylmaleimide. The amino acid sequence of human RNase H1 displays strong  
CC homology with RNase H1 from yeast, chicken, E. coli and mouse. This  
CC sequence was generated from the wild-type human RNase H1 sequence  
CC (ABB83371) by site-directed mutagenesis. It was found that although the  
CC conserved amino acid residues of the putative catalytic site and basic  
CC substrate-binding domain are required for RNase H activity, deletion of  
CC either the catalytic site or the basic substrate-binding domain did not  
CC ablate binding to the heteroduplex substrate. This mutant exhibited an  
CC undetectable rate of cleavage of a heteroduplex substrate, i.e. the  
CC mutation ablated the cleavage activity of the enzyme. Note: The present  
CC sequence was not shown in the specification, but was derived from the  
CC wild-type human RNase H1 sequence shown on Fig 1

XX Sequence 286 AA;

SQ

Query Match 97.5%; Score 1508; DB 5; Length 286;  
Best Local Similarity 97.9%; Pred. No. 1.1e-144;  
Matches 280; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSWLFARVVALALPCRRSGRGFMFYAVRGRKTVFLTMNCRQAQVDRFPAPARFKK 60  
DB 1 MSWLFARVVALALPCRRSGRGFMFYAVRGRKTVFLTMNCRQAQVDRFPAPARFKK 60

QY 61 FATEDENAFVFKKSPVSESGENHQGESPAKCKRLREPLDDGHSAPYAKHMKP 120  
DB 61 FATEDENAFVFKKSPVSESGENHQGESPAKCKRLREPLDDGHSAPYAKHMKP 120

QY 121 SVEPAPVSRDFTFSTMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNGLRPGQ 180  
DB 121 SVEPAPVSRDFTFSTMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNGLRPGQ 180

QY 181 TNQRAEIHAAKCAIOAKTONINKLVLYTDSMFTINGITNNVQGGKMGKTSAGKEVIN 240  
DB 181 TNQRAEIHAAKCAIOAKTONINKLVLYTDSMFTINGITNNVQGGKMGKTSAGKEVIN 240

QY 241 KEDFVALERLTQGMIDQWMTVPVGHSGFIGNEBADRLARAGAQSD 286  
DB 241 KEDFVALERLTQGMIDQWMTVPVGHSGFIGNEBADRLARAGAQSD 286

RESULT 13  
ABB83372 standard; protein; 286 AA.

XX ABB83372;

DT 04-SEP-2002 (first entry)

XX Mutant RNase H1, K226A/K227A.

DE Human; RNase H1; enzyme; mutant.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT 1..73 /label=Region\_I

FT /note="Contains the double-stranded RNA (dsRNA) -binding motif"

FT 74..135 /label=Region\_II

FT 136..286 /label=Region\_III

FT /note="Included within this region are the conserved amino acid residues that form the catalytic site, the divalent cation binding site, and the basic substrate-binding domain"

FT Misc-difference 226 /note="Lys substituted by Ala"

FT Misc-difference 227 /note="Lys substituted by Ala"

FT WO200240635-A2.

XX 23-MAY-2002.

XX 14-NOV-2001; 2001WO-US043929.

XX 15-NOV-2000; 2000US-0248950P.

XX (ISIS-) ISIS PHARM INC.

XX Wu H, Lima WF, Crooke ST;

XX WPI, 2002-519372/55.

DR Novel human RNase H1 polypeptide comprising mutations compared to wild

PT type human RNase H1 useful for inhibiting expression of selected protein

PT by antisense oligonucleotide targeted to RNA encoding selected protein.

XX Claim 5; Page: 39pp; English.

XX The present sequence is a mutant RNase H1 sequence. RNase H hydrolyses  
CC RNA in RNA-DNA hybrids, and the activity of RNase H1 is Mg<sup>2+</sup> dependent  
CC and inhibited by both Mn<sup>2+</sup> and the sulphhydryl blocking agent N-  
CC ethylmaleimide. The amino acid sequence of human RNase H1 displays strong  
CC homology with RNase H1 from yeast, chicken, E. coli and mouse. This  
CC sequence was generated from the wild-type human RNase H1 sequence  
CC (ABB83371) by site-directed mutagenesis. It was found that although the  
CC conserved amino acid residues of the putative catalytic site and basic  
CC substrate-binding domain are required for RNase H activity, deletion of  
CC either the catalytic site or the basic substrate-binding domain did not  
CC ablate binding to the heteroduplex substrate. This mutant exhibited an  
CC initial cleavage rate, of a heteroduplex substrate, of almost two orders  
CC of magnitude slower than the rate observed for the wild-type enzyme.  
CC Note: The present sequence was not shown in the specification, but was  
CC derived from the wild-type human RNase H1 sequence shown on Fig 1

XX Sequence 286 AA;

SQ

Query Match 97.1%; Score 1501; DB 5; Length 286;  
Best Local Similarity 97.6%; Pred. No. 5.8e-144;  
Matches 279; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSWLFARVVALALPCRRSGRGFMFYAVRGRKTVFLTMNCRQAQVDRFPAPARFKK 60  
DB 1 MSWLFARVVALALPCRRSGRGFMFYAVRGRKTVFLTMNCRQAQVDRFPAPARFKK 60

QY 61 FATEDENAFVFKKSPVSESGENHQGESPAKCKRLREPLDDGHSAPYAKHMKP 120  
DB 61 FATEDENAFVFKKSPVSESGENHQGESPAKCKRLREPLDDGHSAPYAKHMKP 120

QY 121 SVEPAPVSRDFTFSTMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNGLRPGQ 180  
DB 121 SVEPAPVSRDFTFSTMGDFVYVYTDGCCSSNGRRPRAGIGYWGPGHPLNGLRPGQ 180

QY 181 TNQRAEIHAAKCAIOAKTONINKLVLYTDSMFTINGITNNVQGGKMGKTSAGKEVIN 240  
DB 181 TNQRAEIHAAKCAIOAKTONINKLVLYTDSMFTINGITNNVQGGKMGKTSAGKEVIN 240

QY 241 KEDFVALERLTQGMIDQWMTVPVGHSGFIGNEBADRLARAGAQSD 286  
DB 241 KEDFVALERLTQGMIDQWMTVPVGHSGFIGNEBADRLARAGAQSD 286

RESULT 14  
ABB83376 standard; protein; 286 AA.

XX ABB83376;

XX 04-SEP-2002 (first entry)  
 XX Mutant RNase H1, K226,227,231,236A.  
 XX Human; RNase H1; enzyme; mutant.  
 XX Homo sapiens.  
 OS Synthetic.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..73  
 FT /label=Region\_I  
 FT /note="Contains the double-stranded RNA (dsRNA) -binding  
 FT motif"  
 FT 74..135  
 FT /label=Region\_II  
 FT 136..286  
 FT /label=Region\_III  
 FT /note="Included within this region are the conserved  
 FT amino acid residues that form the catalytic site, the  
 FT divalent cation binding site, and the basic substrate-  
 FT binding domain"  
 FT Misc-difference 226  
 FT /note="Lys substituted by Ala"  
 FT Misc-difference 227  
 FT /note="Lys substituted by Ala"  
 FT Misc-difference 231  
 FT /note="Lys substituted by Ala"  
 FT Misc-difference 236  
 FT /note="Lys substituted by Ala"  
 PN WO200240635-A2.  
 XX 23-MAY-2002.  
 PD 14-NOV-2001; 2001MO-US043929.  
 XX 15-NOV-2000; 2000US-0248950P.  
 XX (ISIS-) ISIS PHARM INC.  
 PA Wu H, Lima WF, Crooke ST;  
 PI WPI; 2002-519372/55.  
 DR Novel human RNase H1 polypeptide comprising mutations compared to wild  
 PT type human RNase H1 useful for inhibiting expression of selected protein  
 PT by antisense oligonucleotide targeted to RNA encoding selected protein.  
 XX Disclosure; Page; 39pp; English.  
 PS The present sequence is a mutant RNase H1 sequence. RNase H hydrolyses  
 XX RNA in RNA-DNA hybrids, and the activity of RNase H is Mg<sup>2+</sup> dependent  
 CC and inhibited by both Mn<sup>2+</sup> and the sulphhydryl blocking agent N-  
 CC ethylmaleimide. The amino acid sequence of human RNase H displays strong  
 CC homology with RNase H1 from yeast, chicken, E. coli and mouse. This  
 CC sequence was generated from the wild-type human RNase H1 sequence  
 CC (AB83371) by site-directed mutagenesis. It was found that although the  
 CC conserved amino acid residues of the putative catalytic site and basic  
 CC substrate-binding domain are required for RNase H activity, deletion of  
 CC either the catalytic site or the basic substrate-binding domain did not  
 CC ablate binding to the heteroduplex substrate. This mutant exhibited an  
 CC undetectable rate of cleavage of a heteroduplex substrate, i.e. the  
 CC mutations ablated the cleavage activity of the enzyme. Note: The present  
 CC sequence was not shown in the specification, but was derived from the  
 CC wild-type human RNase H1 sequence shown on Fig 1  
 XX  
 SQ Sequence 286 AA;

Query Match 96.3%; Score 1489; DB 5; Length 286;  
 Best Local Similarity 96.9%; Pred. No. 9.7e-143;  
 Matches 277; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSWLFLEAHRVALAALPCRGRSGFGMFYAVRRGRKTGVFLTNNECRAQYDRPPARFKK 60  
 DB 1 MSWLFLEAHRVALAALPCRGRSGFGMFYAVRRGRKTGVFLTNNECRAQYDRPPARFKK 60  
 QY 61 FATEDAMAFVRSASPEVSEGHENHGOSEAKPGKRLREPLDGGHESAOYAGMKP 120  
 DB 61 FATEDAMAFVRSASPEVSEGHENHGOSEAKPGKRLREPLDGGHESAOYAGMKP 120  
 QY 121 SVEPAPVSRDTPSWGDFVWVVTDDCCSSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 DB 121 SVEPAPVSRDTPSWGDFVWVVTDDCCSSNGRRKPRAGIGVYWGPHPLNVGIRLPGRQ 180  
 QY 181 TNGRAEIHAAKCAIEQAKTONINKLVLYTDSMTFINGITNWVGWKKGWKTSAKEVIN 240  
 DB 181 TNGRAEIHAAKCAIEQAKTONINKLVLYTDSMTFINGITNWVGQANGWATSAGAEVIN 240  
 QY 241 KEDFVALERLTQGMDIQMHVPHSGFTIGNEEDRLAREGAKSED 286  
 DB 241 KEDFVALERLTQGMDIQMHVPHSGFTIGNEEDRLAREGAKSED 286  
 RESULT 15  
 ID ADC39108 standard; protein, 269 AA.  
 XX  
 AC ADC39108;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Novel human NOXV polypeptide SEQ ID NO: 50.  
 XX  
 XX antidiabetic; cyostatic; immunomodulator; anorectic; antilipemic;  
 KW neurotropic; neuroprotective; immunostimulant; antiparkinsonian; anti-HIV;  
 KW antiasthmatic; antiinflammatory; hypotensive; antiarteriosclerotic;  
 KW hemostatic; osteopathic; gene therapy; NOXV; diabetes; obesity; cancer;  
 KW lymphoma; uterus cancer; prostate cancer; dyslipidemia; anorexia;  
 KW wasting disorder; Alzheimer's disease; Parkinson's disorder; cachexia;  
 KW cardiomyopathy; AIDS; asthma; Crohn's disease; multiple sclerosis;  
 KW hypertension; atherosclerosis; hemophilia; graft-versus-host disease;  
 KW Albright hereditary osteodystrophy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003010327-A2.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 02-MAY-2002; 2002MO-US014199.  
 XX  
 XX 02-MAY-2001; 2001US-0288063P.  
 PR 03-MAY-2001; 2001US-0288395P.  
 PR 07-MAY-2001; 2001US-0289087P.  
 PR 09-MAY-2001; 2001US-0289817P.  
 PR 09-MAY-2001; 2001US-0289818P.  
 PR 11-MAY-2001; 2001US-0290194P.  
 PR 14-MAY-2001; 2001US-0290753P.  
 PR 15-MAY-2001; 2001US-0291181P.  
 PR 16-MAY-2001; 2001US-0291243P.  
 PR 18-MAY-2001; 2001US-0292001P.  
 PR 21-MAY-2001; 2001US-0292374P.  
 PR 22-MAY-2001; 2001US-0292587P.  
 PR 23-MAY-2001; 2001US-0293107P.  
 PR 25-MAY-2001; 2001US-0293747P.  
 PR 29-MAY-2001; 2001US-0294109P.  
 PR 29-MAY-2001; 2001US-0294110P.  
 PR 30-MAY-2001; 2001US-0294434P.  
 PR 31-MAY-2001; 2001US-0294827P.  
 PR 12-JUL-2001; 2001US-0304879P.  
 PR 31-JUL-2001; 2001US-0308901P.  
 PR 14-AUG-2001; 2001US-0312370P.  
 PR 17-AUG-2001; 2001US-0313416P.  
 PR 10-SEP-2001; 2001US-0318463P.

PR 27-SEP-2001; 2001US-0325683P.  
PR 18-OCT-2001; 2001US-0330292P.  
PR 28-NOV-2001; 2001US-033873P.  
PR 03-DEC-2001; 2001US-0338909P.  
PR 03-DEC-2001; 2001US-0337552P.  
PR 21-FEB-2002; 2002US-0359245P.  
PR 01-MAY-2002; 2002US-00136826.

XX (CURA-) CURAGEN CORP.

PI Miller CE, Kekuda R, Malpankar UM, Li L, Pena CE, Spytek KA,  
PI Gorman L, Guo X, Fernandes ER, Smithson G, Stone DJ, Zeehuysen BD,  
PI Paturajan M, Anderson DW, Mezei PS, Peyman JA, Macdougall JR,  
PI Padigar M, Raselli L, Shenoy SG, Gerlach VL, Shinkens RA, Zhong M,  
PI Edinger SR, Ellerman K;  
XX WPI: 2003-239445/23.  
DR N-PSDB; ADC39107.

PT New NOVX polypeptides and polynucleotides, useful in gene therapy,  
PT particularly for treating or preventing a syndrome associated with a  
PT human disease e.g. diabetes, obesity, cancer, Alzheimer's disease,  
PT hypertension or hemophilia.

PS Claim 1; SEQ ID NO 50; 748bp; English.

XX The invention relates to new isolated NOVX polypeptides, the genes  
XX encoding them or sequences having at least 95% identity to the amino acid  
XX or nucleotide sequences. The NOVX polypeptide is useful as a therapeutic,  
XX particularly in the manufacture of a medicament for treating a syndrome  
XX associated with a human disease, which includes a pathology associated  
XX with NOVX polypeptide. The NOVX polypeptide is particularly useful for  
XX treating, preventing or alleviating pathology associated with NOVX  
XX polypeptide in a mammal, e.g. a human. The NOVX nucleic acid and  
XX polypeptide are especially useful for treating or preventing e.g.  
XX diabetes, obesity, cancer (e.g. lymphoma, uterine cancer or prostate  
XX cancer), dyslipidemias, anorexia, wasting disorders, Alzheimer's disease,  
XX Parkinson's disorder, cachexia, cardiomyopathy, AIDS, asthma, Crohn's  
XX disease, multiple sclerosis, hypertension, atherosclerosis, hemophilia,  
XX graft-versus-host disease or Albright hereditary osteodystrophy. The DNA  
XX encoding the protein is useful in gene therapy for treating the above  
XX conditions. These are also useful in developing powerful assay system for  
XX functional analysis of various human disorders, as well as in diagnostic  
XX applications. This sequence represents one of the NOVX proteins of the  
XX invention.

CC Sequence 269 AA;

Query Match 78.8%; Score 1218; DB 7; Length 269;  
Best Local Similarity 88.0%; Pred. No. 3.4e-115;  
Matches 235; Conservative 7; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRGRKTVFLTWNECRAOVDFPAPRFXK 60  
DB 1 MSRLSLAARVALAALPCRR-SRGFMFYAVRGRKSAVFLTNGNECKAQVDFPAPRFXK 59  
QY 61 FATEDEANAFVFKSAPVSESGENHQESAEKPKRLREPLDGDGHSAPYAKHMP 120  
DB 60 FATEDEANDFVKASPEVSEGOENHQESSTKSKRLREPLDGDGDSASPYASTMP 119  
QY 121 SYEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYVWGPGLNGLRPGQ 180  
DB 120 SYKPAVPSRDTFTSYMGDFVYVYTDGCCSSNGRRKPRAGIRYWGPGYFLYTGIRLPGQ 179  
QY 181 TNQRAEIHAAKCAIAEQATONINKLVLYTDSMFTINGITNNVQGWKNGMKTSAKEVIN 240  
DB 180 TNQRAEIHAAKCAIEAQATOKINKLVLYTDSMFTINGITNNVQGWKNGMKTSAKEVIN 239  
QY 241 KEDFVALERLTQGMIDQW-----MAYP 262  
DB 240 KEDFVALERLTQGMIDQWASINVAHP 266

RESULT 16  
ABP53041  
ID ABP53041 standard; protein; 285 AA.

XX AC ABP53041;

DT 06-NOV-2002 (first entry)

DE Mouse RNase H1 protein sequence SEQ ID NO:11.

XX RNase H, antisense technology; inhibition.

OS Mus sp.

PN WO200264841-A1.

PD 22-AUG-2002.

PF 12-FEB-2002; 2002WO-US004243.

PR 12-FEB-2001; 2001US-00781712.

PA (ISIS-) ISIS PHARM INC.

PI Crooke ST, Lima WF, Wu H;

XX WPI: 2002-657606/70.

PT Use of a mammalian, particularly human, RNase H, for treating an animal  
PT with a disease or condition associated with a human RNase H, for  
PT inhibiting the expression of a protein, or for reducing cellular RNA via  
PT antisense technology.

PS Claim 8; Page 60-61; 70pp; English.

XX The present invention describes a method for promoting the inhibition of  
XX the expression of a protein comprising employing a mammalian RNase H  
XX polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA  
XX complex duplex occurs. Also described is a compound 8 to 50 nucleobases  
XX in length targeted to the nucleic acid encoding the human RNase H1  
XX polypeptide, where the compound specifically hybridizes with and inhibits  
XX the expression of a human RNase H1 polypeptide. The compound, which is  
XX an antisense oligonucleotide, is useful for inhibiting the expression of  
XX a human RNase H1 polypeptide in cells or tissues, as well as for  
XX treating an animal with a disease or condition associated with a human  
XX RNase H1 polypeptide. The method is useful for inhibiting the expression  
XX of a protein, particularly for reducing cellular RNA via antisense  
XX technology. The present sequence represents a mouse RNase H1 protein  
XX sequence, given in the exemplification of the present invention

CC Sequence 285 AA;

Query Match 77.5%; Score 1197.5; DB 5; Length 285;  
Best Local Similarity 77.3%; Pred. No. 4.5e-113;  
Matches 221; Conservative 27; Mismatches 37; Indels 1; Gaps 1;

QY 1 MSWLLFLAHRVALAALPCRRGSRGFMFYAVRGRKTVFLTWNECRAOVDFPAPRFXK 60  
DB 1 MSRLSLAARVALAALPCRR-SRGFMFYAVRGRKSAVFLTNGNECKAQVDFPAPRFXK 59  
QY 61 FATEDEANAFVFKSAPVSESGENHQESAEKPKRLREPLDGDGHSAPYAKHMP 120  
DB 60 FATEDEANDFVKASPEVSEGOENHQESSTKSKRLREPLDGDGDSASPYASTMP 119  
QY 121 SYEPAPVSRDFTSYMGDFVYVYTDGCCSSNGRRKPRAGIGYVWGPGLNGLRPGQ 180  
DB 120 SYKPAVPSRDTFTSYMGDFVYVYTDGCCSSNGRRKPRAGIRYWGPGYFLYTGIRLPGQ 179  
QY 181 TNQRAEIHAAKCAIAEQATONINKLVLYTDSMFTINGITNNVQGWKNGMKTSAKEVIN 240  
DB 180 TNQRAEIHAAKCAIEAQATOKINKLVLYTDSMFTINGITNNVQGWKNGMKTSAKEVIN 239  
QY 241 KEDFVALERLTQGMIDQWMEVPGHSGFIGNEADRLAREGAKQSD 286



DB 240 KEDFVLELITQGMIDQMHIPGHSFVGNBEADRLAREGAKQSED 285

RESULT 17  
ADA05718  
ID ADA05718 standard; protein; 203 AA.

AC ADA05718;

DT 06-NOV-2003 (first entry)

DE Human NOV15a protein SEQ ID NO:78.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
XX immunomodulator; cytoskeletal; neurotrophic; neuroprotective;  
XX antiparkinsonian; antihypertensive; gene therapy; human disease;  
XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

PN WC02003029424-A2.

PD 10-APR-2003.

PF 02-OCT-2002; 2002WC-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 12-OCT-2001; 2001US-0328056P.

PR 15-OCT-2001; 2001US-0328449P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 22-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 17-MAY-2002; 2002US-0381042P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383811P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

(CURA-) CURAGEN CORP.

PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Ratturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerkusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Raschell L, Stone DJ, Pena CE, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergs C, Didiyko VA;  
PI Eisen AJ, Gangoli EA, Rieger DK, Spaderna SK;

WPI; 2003-381626/36.

DR N-PSDB; ADA05717.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidaemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

PS Claim 1; Page 159; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55  
XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
XX described above and a carrier; (2) a kit comprising, in one or more  
XX containers, the composition described above; (3) an isolated nucleic acid  
XX molecule which encodes a NOVX protein of the invention; (4) a vector  
XX comprising the nucleic acid molecule described above; (5) a cell  
XX comprising the above vector; (6) an antibody that immunospecifically  
XX binds to the polypeptide described above; (7) methods for determining the  
XX presence or amount of the above polypeptide or nucleic acid molecule in a  
XX sample; (8) methods for determining the presence of or predisposition to  
XX a disease associated with altered levels of expression of the above  
XX polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
XX method of identifying an agent that binds to the polypeptide described  
XX above; (10) a method for identifying a potential therapeutic agent for  
XX use in treating a pathology that is related to an aberrant expression or  
XX aberrant physiological interactions of the polypeptide; (11) a method of  
XX screening for a modulator of activity or of latency or predisposition to  
XX a pathology associated with the polypeptide; (12) a method for modulating  
XX the activity of the polypeptide described above; (13) methods of treating  
XX or preventing a pathology associated with the above polypeptide in a  
XX mammal; and (14) a method for producing the above polypeptide. NOVX  
XX sequences have antidiabetic, anorectic, antibacterial, virucide,  
XX immunomodulator, cytoskeletal, neurotrophic, neuroprotective, antiparkinsonian  
XX and antihypertensive activities, and can be used in gene therapy. The  
XX polypeptide is useful in manufacturing a medicament for treating a  
XX syndrome associated with a human disease. The polypeptide or the nucleic  
XX acid molecule may be used to diagnose, treat or prevent metabolic  
XX disorders such as diabetes or obesity, infections, cachexia, cancer,  
XX neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
XX disease, immune disorders, haematopoietic disorders and various  
XX dyslipidaemias. The nucleic acids can also be used as hybridisation  
XX probes, in chromosome mapping, tissue typing, preventive medicine and  
XX pharmacogenomics. The present sequence represents a human NOVX from the  
XX present invention.

SO Sequence 203 AA;

Query Match 67.6%; Score 1045.5; DB 6; Length 203;  
Best Local Similarity 70.3%; Pred. No. 8.3e-98;

Matches 201; Conservative 1; Mismatches 1; Indels 83; Gaps 1;

QY 1 MSWLFLLAHRVVALPCRRGSGFGMFYAVRGRTGVLTNBERAOVDRPAPRPFK 60

DB 1 MSWFLFLAHRVVALPCRRGSGFGMFYAVRGRTGVLTNBERC----- 46

QY 61 FATEDEAWAFVRKASPEVSEGHENHGOSEAKPGKRLREPLDGDGHSAQPYAKMKP 120

DB 47 ----- 46

QY 121 SVEPAPVSRDRTSYNGDFVVTYTTDCCSSNGRRKRPAGIGYWGQPLNVLPGRO 180

DB 47 -----RDTFSYNGDFVVTYTTDCCSSNGRRRPRAGIGYWGQPLNVLPGRO 97

QY 181 TNQRAETHAACKIAEQAKTONIKLVLYTDSMTTIGTINWVGKMGKTSAGKEVIN 240

DB 98 TNQRAETHAACKIAEQAKTONIKLVLYTDSMTTIGTINWVGKMGKTSAGKEVIN 157

QY 241 KEDFVLELITQGMIDQMHIPGHSFVGNBEADRLAREGAKQSED 286

DB 158 KEDFVLELITQGMIDQMHIPGHSFVGNBEADRLAREGAKQSED 203

RESULT 18

ADA05720  
ID ADA05720 standard; protein; 210 AA.

AC ADA05720;



XX 06-NOV-2003 (first entry)  
 DT Human NOV15b protein SEQ ID NO:80.  
 DE  
 XX human; NOX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 OS  
 XX MO2003029424-A2.  
 XX  
 XX 10-APR-2003.  
 PD  
 XX  
 XX 02-OCT-2002; 2002MO-US031373.  
 PF  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0345757P.  
 PR 01-NOV-2001; 2001US-0346376P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Paturujan M, Spytek KA, Edinger SR, Ellerman K, Malvanekar UM;  
 PI Ort T, Gorman L, Zerkowen BD, Anderson DW, Zhong M, Cateron E;  
 PI Ji W, Miller CB, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shinkens R, Rothenberg ME, Leach MD, Ague MW, Bergs C, Dippio VA;  
 PI Eisen AJ, Gangoli EA, Rieger DK, Spaderna SK;  
 XX  
 XX MPI; 2003-381626/36.  
 DR N-PSDB; ADA05719.  
 DR  
 XX  
 XX New NOX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidaemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 XX Claim 1; Page 159; 586pp; English.  
 PS  
 XX The present invention describes NOX proteins, where X can be 1 to 55  
 CC (e.g. NOX1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antiparkinsonian activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOX from the  
 CC present invention.  
 XX  
 SQ Sequence 210 AA;

Query Match 67.6%; Score 1045.5; DB 6; Length 210;  
 Best Local Similarity 70.3%; Pred. No. 8.7e-98;  
 Matches 201; Conservative 1; Mismatches 83; Gaps 1;

QY 1 MSWFLFLHRAVALALPCRRSGRGFMFAYARRGKTVPLTWNCRAQVDFPAPARFX 60  
 DB 5 MSWFLFLHRAVALALPCRRSGRGFMFAYARRGKTVPLTWNC----- 50  
 QY 61 FATEDAWAFVRSKASPEVSEGHENQGESEAKGKRLREPLDGDHESAPYAKHMKP 120  
 DB 51 ----- 50  
 QY 121 SYEPAPVSRDTFSYMGDFVYVYTDGCCSSNGRRKRRPAGIGYWGPGHPLNVGIRLPGRQ 180  
 DB 51 -----RDTFSYMGDFVYVYTDGCCSSNGRRRPRAGIGYWGPGHPLNVGIRLPGRQ 101  
 QY 181 TNGRAEHAACALIQATQNTINKLVYTDNSFTNGTNNVQGMKRWKTSACKREVIN 240  
 DB 102 TNGRAEHAACALIQATQNTINKLVYTDNSFTNGTNNVQGMKRWKTSACKREVIN 161  
 QY 241 KEDFVALERLTQGMIDQMWVHGSGFIGNERADRLAREGAKQSD 286  
 DB 162 KEDFVALERLTQGMIDQMWVHGSGFIGNERADRLAREGAKQSD 207

RESULT 19  
 ADA05722  
 ID ADA05722 standard; protein; 195 AA.  
 XX  
 AC ADA05722;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 XX Human NOV15c protein SEQ ID NO:82.  
 XX  
 XX human; NOX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia.  
 XX Homo sapiens.  
 OS  
 PN WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 PF 02-OCT-2002; 2002MO-US031173.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 09-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373864P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Fatturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ott T, Gorman L, Zetluseen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rasetelli L, Stone DJ, Pena CEA, Shenoy SG;  
 PI Shimetsu RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiPippo VA,  
 PI Eisen AV, Gangoli EA, Rieger DK, Spaderma SK;  
 XX  
 XX WPI; 2003-381626/36.  
 DR N-PSDB; ADA05721.  
 XX  
 PT New NOXV polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOXV-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 160; 586pp; English.  
 XX  
 CC The present invention describes NOXV proteins, where X can be 1 to 55  
 CC (e.g. NOXV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOXV protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOXV  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antihypertensive activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridization  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOXV from the  
 CC present invention.  
 XX

SO Sequence 195 AA:

Query Match 63.4%; Score 979.5; DB 6; Length 195;  
 Best Local Similarity 68.9%; Pred. No. 4.1e-91;  
 Matches 188; Conservative 2; Mismatches 0; Indels 83; Gaps 1;

QY 14 AALPCRGRSGRMFPAVRGRGTGVFLTMNEGRQAVDRPARFKFATEDAMAFVRK 73  
 DB 3 SALPCRGRSGRMFPAVRGRGTGVFLTMNEC----- 35  
 QY 74 SASPEVSEGHENQGESEAKPGKRLREPLDGDGHSAPYAKMKRSPVAPVSRDTF 133  
 DB 36 -----RDTF 39  
 QY 134 SYNGDFVYVYTTDCCSSNGRRKPRRAGIGYWGFGHPLANGIRLPGQTNRRAIHAACRA 193  
 DB 40 SYMGDFVYVYTTDCCSSNGRRRPRRAGIGYWGFGHPLANGIRLPGQTNRRAIHAACRA 99  
 QY 194 IEQAKTONIKLYLYTDSMTTITGTTNWQGMKNGMTSAGKEVYNKEPVALERTTG 253  
 DB 100 IEQAKTONIKLYLYTDSMTTITGTTNWQGMKNGMTSAGKEVYNKEPVALERTTG 159  
 QY 254 MDIQMHVPGHSGFIGNESADRLAREGAKQSED 286  
 DB 160 MDIQMHVPGHSGFIGNESADRLAREGAKQSED 192

RESULT 20

AAV25098  
 ID AAV25098 standard; protein; 216 AA.

AC AAV25098;

DT 24-AUG-1999 (first entry)

DE Mouse RNase H homologue protein fragment.

KW Type 2 RNase H; oligonucleotide-RNA duplex; cleavage; antisense therapy;  
 KW interaction; target mRNA; mouse.

OS Mus sp.

PN WO9928447-A1.

PD 10-JUN-1999.

PF 02-DEC-1998; 98MO-US025488.

PR 04-DEC-1997; 97US-0067458P.

PA (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 1999-394857/33.

DR New isolated human RNase H polypeptide and nucleic acid, used to develop  
PT agents for enhancing the efficiency of antisense therapy.

XX Discloure; Page 33-34; 37pp; English.

XX This invention describes a novel human RNase H polypeptide and nucleic  
CC acid which can be used to cleave the RNA strand of oligonucleotide-RNA  
CC duplexes and to develop agents for enhancing the efficiency of antisense  
CC therapy. The products of the invention are useful in defining the  
CC interaction of human type 2 RNase H and antisense oligonucleotides and  
CC identifying methods for enhancing this interaction so that antisense  
CC oligonucleotides are more effective at inhibiting their target mRNA. The  
CC products can be used for enhancing the efficacy of antisense  
CC oligonucleotide therapies. This sequence represents a mouse RNase H  
CC homologue described in the method of the invention

XX Sequence 216 AA;

Query Match 62.0%; Score 959; DB 2; Length 216;

Best Local Similarity 70.2%; Pred. No. 5.8e-89;

Matches 177; Conservative 20; Mismatches 19; Indels 36; Gaps 1;

QY 21 GSRGFMFYAVRGRKTVFLTWNECRAQVDRFPARFKKATDEDAVFKASPEVS 80

DB 1 GICGGMFYAVRGRRPGVFLSWSECKAQVDRFPARFKKATDEDAVFKASPEVS 60

QY 81 EGHENQHGESEAKKRGKRLREPLDDGHSAPYAKMKPSVEPAPVSRDFTSYMGDFV 140

DB 61 KOESAHHEKKSQAKTSKRPREPL-----V 84

QY 141 VYTTGGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRTNORAEIHAACKAIEQAKTQ 200

DB 85 VYTTGGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRTNORAEIHAACKAIEQAKTQ 144

QY 201 NINKLVLYTDSMFTINGITNWTWQGWKKGWKTSAKEVINKEDEPVALERTLQGMIDQW 260

DB 145 NISKLVLYTDSMFTINGITNWTWQGWKKGWKTSAKEVINKEDEPVALERTLQGMIDQW 204

QY 261 VPGHSGFTIGNEE 272

DB 205 IPGHSGFVGNNEE 216

RESULT 21

AAB97511 standard; protein; 216 AA.

XX AAB97511;

DT 14-AUG-2001 (first entry)

XX E coli type II RNase H protein.

XX E coli; RNase H type II; RNase H1 cleavage substrate; antisense therapy;

KW gene therapy.

OS Escherichia coli.

XX WO200123613-A1.

PD 05-APR-2001.

PF 29-SEP-2000; 2000MO-US026729.

PR 30-SEP-1999; 99US-00409926.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H, Manoharan M;

XX WPI; 2001-343164/36.

DR Chimeric oligonucleotides that can serve as substrates for human RNase  
PT H1, useful for enhancing the effectiveness of antisense gene therapies.

XX Discloure; Fig 6; 178pp; English.

XX The present invention provides a number of DNA-RNA oligonucleotides which  
CC can act as substrates for human RNase H1 (a type II RNase). The sequence  
CC consists of two portions, one of which is capable of supporting cleavage  
CC of a complementary target RNA and the other of which is incapable of  
CC supporting such cleavage. These can be used to enhance the effectiveness  
CC of antisense therapies. The present sequence is E. coli RNase H1

XX Sequence 216 AA;

Query Match 62.0%; Score 959; DB 4; Length 216;

Best Local Similarity 70.2%; Pred. No. 5.8e-89;

Matches 177; Conservative 20; Mismatches 19; Indels 36; Gaps 1;

QY 21 GSRGFMFYAVRGRKTVFLTWNECRAQVDRFPARFKKATDEDAVFKASPEVS 80

DB 1 GICGGMFYAVRGRRPGVFLSWSECKAQVDRFPARFKKATDEDAVFKASPEVS 60

QY 81 EGHENQHGESEAKKRGKRLREPLDDGHSAPYAKMKPSVEPAPVSRDFTSYMGDFV 140

DB 61 KOESAHHEKKSQAKTSKRPREPL-----V 84

QY 141 VYTTGGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRTNORAEIHAACKAIEQAKTQ 200

DB 85 VYTTGGCCSSNGRRPRAGIGYWGPGHPLNVGIRLPGRTNORAEIHAACKAIEQAKTQ 144

QY 201 NINKLVLYTDSMFTINGITNWTWQGWKKGWKTSAKEVINKEDEPVALERTLQGMIDQW 260

DB 145 NISKLVLYTDSMFTINGITNWTWQGWKKGWKTSAKEVINKEDEPVALERTLQGMIDQW 204

QY 261 VPGHSGFTIGNEE 272

DB 205 IPGHSGFVGNNEE 216

RESULT 22

AAB97509 standard; protein; 293 AA.

XX AAB97509;

DT 14-AUG-2001 (first entry)

XX Chicken type II RNase H protein.

XX Chicken; RNase H type II; RNase H1 cleavage substrate; antisense therapy;

KW gene therapy.

OS Gallus gallus.

XX WO200123613-A1.

PD 05-APR-2001.

PF 29-SEP-2000; 2000MO-US026729.

PR 30-SEP-1999; 99US-00409926.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H, Manoharan M;

XX WPI; 2001-343164/36.



PR 18-OCT-2001; 2001US-0330303P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0344629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 16-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyanar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DM, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;  
 PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Bergins C, Dipippo VA;  
 PI Eileen AJ, Gangoli EA, Rieger DK, Spaderma SK;  
 XX  
 XX MPI; 2003-381626/36.  
 DR N-PSDB; ADA05723.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics.  
 XX  
 PS Claim 1; Page 160; 586pp; English.

CC present invention.  
 XX  
 SQ Sequence 152 AA;  
 Query Match 51.2%; Score 792; DB 6; Length 152;  
 Best Local Similarity 99.3%; Pred. No. 3.5e-72;  
 Matches 145; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 137 GDFVVVYTDCCSSNGRRKPRAGIVVWGPQHPLNVGIRLPGRTNORABITHACKAIQ 196  
 DB 4 GDFVVVYTDCCSSNGRRRPRAGIVVWGPQHPLNVGIRLPGRTNORABITHACKAIQ 63  
 QY 197 AKTONINKLVLYTDSMTFTINGITVWQGMKNGMTSAGKEVINKEDEVALERLTQGMNDI 256  
 DB 64 AKTONINKLVLYTDSMTFTINGITVWQGMKNGMTSAGKEVINKEDEVALERLTQGMNDI 123  
 QY 257 QMHHVPHSGFTIGNEADRLAREGAK 282  
 DB 124 QMHHVPHSGFTIGNEADRLAREGAK 149  
 RESULT 25  
 ABB58440  
 ID ABB58440 standard; protein; 333 AA.  
 XX  
 AC ABB58440;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 2112.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-0064150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR MPI; 2001-656860/75.  
 DR N-PSDB; ABL02543.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 2112; 21bp + Sequence listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signaling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
 CC ABB12072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WFO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 333 AA;  
 Query Match 27.3%; Score 422; DB 4; Length 333;  
 Best Local Similarity 34.4%; Pred. No. 4.8e-34;

